# DOCUMENT RESUME

ED 236 863 EC 160 884

| | |
|---|---|
| AUTHOR | Oakland, Thomas, Ed. |
| TITLE | Nonbiased Assessment. |
| INSTITUTION | Minnesota Univ., Minneapolis. National School Psychology Inservice Training Network. |
| SPONS AGENCY | Office of Special Education and Rehabilitative Services (ED), Washington, DC. |
| PUB DATE | 81 |
| GRANT | G00784657 |
| NOTE | 160p.; For related documents, see EC 160 883 and EC 160 885. |
| AVAILABLE FROM | National School Psychology Inservice Training Network, Psychology in the School Program, N532 Elliott Hall, 75 E. River Rd., Minneapolis, MN 55455 ($14.50). |
| PUB TYPE | Guides - Non-Classroom Use (055) |
| EDRS PRICE | MF01 Plus Postage. PC Not Available from EDRS. |
| DESCRIPTORS | *Culture Fair Tests; *Disabilities; Elementary Secondary Education; Emotional Disturbances; Intervention; Language Tests; *Limited English Speaking; Mild Mental Retardation; *Minority Groups; Models; *Student Evaluation; Test Bias; Testing Problems |

ABSTRACT

The Nonbiased Assessment module contains seven sections that provide information on various issues relevant to the construction of programs for assessing minority group children. Each section, in addition to the discussion of the topic, contains a list of objectives for the workshop participant or reader, a pretest, simulations, and an annotated bibliography of recommended readings. Section I, on basic considerations regarding nonbiased assessment (by T. Oakland), covers components of a diagnostic intervention process; biasing factors (parent influences, child characteristics, examiner characteristics, diagnostic-intervention techniques, and school system policies and practices); and five diagnostic intervention models (medical, social system, psychoeducational process, educational task analysis, and pluralistic). A second section by D. Bersoff reviews legal principles with reference to P.L. 94-142 (the Education for All Handicapped Children Act) and the Rehabilitation Act of 1973. Sociocultural considerations are addressed by R. Henderson in Section III. In Section IV, J. Matluck and B. Mace-Matluck consider the nature of language, language assessment characteristics of oral language proficiency tests, legally mandated assessment and intervention, and a diagnosis of learning difficulties in limited English speaking and bilingual children. Section V (by A. Hofmeister and C. Preston) gives an overview of educational assessment and its purposes. A sixth section, by D. Reschly, focuses on appropriate assessment for mildly retarded children; while a final section, by M. Tombari, describes nonbiased assessment of emotionally disturbed students. (SW)

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED236863

# NONBIASED ASSESSMENT

Thomas Oakland
Module Editor
*The University of Texas at Austin*

EC160884

A project of The National School Psychology Inservice Training Network
University of Minnesota, Minneapolis, Minnesota

The project presented herein was supported by Grant No. G00784657 from the Office of Special Education to the University of Minnesota. The points of view expressed in this publication are those of the authors and do not necessarily reflect the positions of the U.S. Department of Education and no official endorsement by the U.S. Department of Education should be inferred.

The material presented herein is one of three modules developed and produced for The National School Psychology Training Network, James Ysseldyke, Director.

Minneapolis, Minnesota
1981

Additional copies of this module may be ordered from the Upper Midwest Regional Resource Center, 2037 University Avenue S.E., Suite 101, University of Minnesota, Minneapolis, Minnesota 55414. Prices for single and multiple copies will be furnished upon request.

# Table of Contents

Foreword ........ iii

Introduction ........ iv

Nonbiased Assessment: Basic Considerations
Thomas Oakland ........ 1

Nonbiased Assessment: Legal Principles
Donald Bersoff ........ 19

Nonbiased Assessment: Sociocultural Considerations
Ronald W. Henderson ........ 35

Nonbiased Assessment: Understanding Language Characteristics
Joseph H. Matluck & Betty J. Mace-Matluck ........ 57

Nonbiased Educational Assessment
Alan M. Hofmeister & Caroline N. Preston ........ 77

Nonbiased Assessment and the Mentally Retarded Student
Daniel Reschly ........ 97

Nonbiased Assessment and the Emotionally Disturbed Student
Martin Tombari ........ 123

# Acknowledgments

The development and production of this module on Nonbiased Assessment occurred through the collective efforts of many persons. Its origins stem from the thoughts and efforts of James Ysseldyke, principal investigator, and Maynard Reynolds and Richard Weinberg, co-investigators of The National School Psychology Inservice Training Network. They, together with Jack Bardon, Ann Engin, Ronald Fishman, Walter Hodges, Leon Hall, Mary Jo MacGregor, and I provided input to the original proposal for this project.

I am most appreciative of the contributions from the writers: Donald Bersoff, Ronald Henderson, Alan Hofmeister, Caroline Preston, Betty and Joseph Matluck, Daniel Reschly, and Martin Tombari. They put aside other important responsibilities to accommodate our frequent requests for assistance: to them the strongest credit is due.

Members of the Texas Advisory Group played a key role by first identifying the content necessary for each section of the module and later by reviewing and constructively evaluating the manuscripts. The members are Juana Canabal Antokoletz, Lynn Frank, and Floyd Hill. Walter Borg also provided helpful ideas on the module's design and evaluation.

The module also was reviewed and constructively evaluated by members of the Task Force for the National School Psychology Inservice Training Network. They include Jack Bardon, Ann Engin, Leon Hall, Walter Hodges, Sylvia Rosenfield, Winfred Tillery, James Tucker, and LaMonte Wyche. The project staff at the University of Texas at Austin (Debbie Boswell Hirsch and Mary Attwell Hartye) and at The University of Minnesota National School Psychology Inservice Training Network (James Moran, Pam Ashley, Joel Meyers, and Maureen Koenen) also deserve special recognition.

Our collective efforts were directed toward one major goal: to develop materials which promote the use of quality methods for assessing minority group children. We hope the materials are found to be commensurate with this goal by a receptive audience.

Thomas Oakland
Module Editor
Austin, Texas

## Foreword

The recent and significant revisions in public policy on the education of handicapped students are reflected in the intent and provisions of Public Law 94-142. Implicit in the Law is the need for profound change in the roles, training, and retraining of all public school personnel, especially teachers and school psychologists, who are directly involved in the placement and educational programming of all pupils with any degree of handicap. Since school psychology is a basic resource for all educators, the National School Psychology Inservice Training Network has been engaged in the development of inservice training materials designed expressly for the purpose of providing training to practicing school psychologists in the use of innovative assessment procedures that will enable them to assist the schools and special and regular education teachers to comply with the provisions of PL 94-142. The Network, funded by the U.S. Office of Special Education, has worked for the past two years as a temporary system in support of the standing structures relevant to the profession of school psychology (i.e., State School Psychology Consultants, the National Association of School Psychologists, the School Psychology Division of the American Psychological Association, the Council of Directors of School Psychology Programs, State and Local School Psychology Associations).

Broadly speaking, the primary responsibility of school psychologists always has been and will continue to be one of working with teachers to facilitate learning by children. In meeting this responsibility, school psychologists bring to education a rich background in the theory and applications of psychology, from the principles derived in experimental laboratories to those derived in clinical settings. Indeed, school psychology is in some sense responsible and accountable for making the principles of educational psychology, developmental psychology, clinical psychology, and general psychology available to school personnel in order to enhance children's learning. Throughout this module, school psychology is conceptualized within this broad context.

School psychologists have had and will continue to have considerable responsibility for the psychoeducational decisions that affect the placement and programming of handicapped children. Currently, approximately 12,000 school psychologists are employed in the nation's schools. Because of the tremendous diversity in the nature and level of training which has been available to them, they demonstrate considerable variance in competency for the functions they are required to carry out under P.L. 94-142. The knowledge base exists to ensure their competency to comply with legislative mandates; to date, however, no systematic national effort had been made to assure that handicapped children are served by highly competent school psychologists.

The variance in competency displayed by school psychologists indicates the lack of a consistent conception of their role and, thus, of preparation. This module, and others prepared by the Network, represent an initial effort to provide a needed training resource.

Special appreciation is extended to the many professionals who have played significant roles in the guidance of the Network and in advising us in our efforts along the way. Many school psychologists have participated in field testing these materials, either at local meetings or at preconvention workshops held at the annual meetings of NASP or APA. Their feedback on initial drafts of these materials contributed to a better final product. Professors Maynard Reynolds and Richard Weinberg at Minnesota, Network co-investigators, contributed significantly to the initial impetus for this effort. Joel Meyers, Professor of School Psychology at Temple University, spent his sabbatical during the 1979-80 academic year as Director of the Network, being assisted in his efforts by Maureen Koenen. Sylvia Rosen and Jacqueline Schakel did the technical editing necessary to produce this final product.

Thomas Oakland, the coordinator of efforts to develop this module, deserves special recognition for identifying those professionals who could best contribute to the development of this module, and for working so very hard to bring to fruition their efforts and produce a product to assist school psychologists in the fair assessment of students.

James E. Ysseldyke, Director
National School Psychology
Inservice Training Network

# Introduction

School psychologists are professionals with intensive training in psychology and extensive training and experiences in education who are committed to providing quality psychoeducational services to students. As professionals in the schools, school psychologists can be expected to provide one-of-a-kind services based on their knowledge of individual differences and the recognition that each student is unique. This awareness that no two students are alike has guided school psychologists in selecting assessment tools and intervention techniques geared to each child's unique situation.

The confidence of school psychologists in their ability to provide quality services, particularly with respect to minority group children, has been shaken recently by allegations, judicial decisions, state and federal mandates and other challenges to their effectiveness. The commitment of most school psychologists to provide appropriate services to students has not been weakened. However, some confusion has arisen over how to work with children and parents who come from different sociocultural and racial ethnic groups.

This module, Nonbiased Assessment, has been prepared in an attempt to answer some of the questions about assessing minority group children that have surfaced from the confusion. Its content was specified by practicing school psychologists and approved by directors of school psychological and special education services from some of our nation's largest school districts. It represents an approach to nonbiased assessment that we believe allows the school psychologist to best meet the needs of individual students while also staying within the guidelines of federal and state regulations. Just as each child we evaluate is unique, so is each school district, school, and school psychologist. Therefore, ideas and materials contained here will have to be adapted to different situations.

## Organization of the Module

The Nonbiased Assessment module contains seven sections that provide information on various issues which are relevant to the construction of nonbiased programs in school systems. In each section, in addition to the discussion of the topic, are a table of contents, a list of objectives for the workshop participant or reader, a pretest, simulations, and an annotated bibliography of recommended readings. In some sections a posttest is also provided. In others the pretest may also serve as a posttest. Within the narrative text of each section are several short tasks or questions which allow the participant to assess his or her understanding of the material as it is discussed.

## Use of the Materials

The materials in the Nonbiased Assessment Module lend themselves to a variety of instructional formats. Because the materials were designed for the purpose of continuing the education of practicing school psychologists, it is assumed that most users of these materials will receive training through a workshop format. If all seven sections are presented in one workshop, a two- or three-day session will be required. However, each of the sections is self-contained and can be presented separately in two or three hours depending on the section. Since some school districts are unable to set aside a 2 to 3 day period of time at once for inservice, the module can be presented in a series of half-day workshops throughout the school year.

In field testing this module we found that participants appreciated the workshop format because it allowed a sharing of ideas among several people. However, materials may also be used for team self-instruction or individual self-instruction. This may be especially productive for individual or pairs of school psychologists in rural or remote areas who do not have frequent access to continuing education opportunities.

In any instructional format it is strongly encouraged that participants complete the pre- and posttests and all tasks and simulations. Although knowledge about how to proceed is important, knowledge put into constructive *actions* is a more powerful tool to build strong, defensible, and effective nonbiased assessment practices.

## Suggestions to Workshop Presenters

A thorough acquaintance with the Nonbiased Assessment materials is a must for the potential workshop presenter. In order to make the workshop as relevant and timely as possible to the participants, it is also crucial that the presenter know the needs of the audience to help in scheduling the workshop, outlining areas of emphasis, and making simulations as helpful as possible. The following suggestions are offered as guidelines in preparing a workshop on nonbiased assessment.

### Scheduling

The workshop presenter will, of course, need to choose a schedule which will best fit the needs of the participants. In some cases a weekend workshop which covers all the materials may be feasible. In other cases practicing school psychologists may find it more convenient to attend a series of shorter sessions throughout the school year. This second option has the advantage of allowing participants to try out new skills and bring back questions and feedback to later sessions. In any case familiarity with the workshop materials, inservice needs of participants, and time constraints is necessary before scheduling sessions.

### Potential Audience

In most cases the presenter's audience will be practicing school psychologists. Use of the Nonbiased Assessment Module need not be limited to school psychologists, however. Special education personnel, other school diagnostic staff, and administrators may also benefit from knowledge in this area. In each different case the materials could be easily adapted.

In field testing these materials we found that school psychologists rated sections as more or less valuable depending on their level of expertise in each section's subject matter. Therefore, a highly sophisticated audience may need less emphasis on certain sections, whereas a less knowledgeable group may require equal emphasis on all sections. If the participants are given the materials to read ahead of time, discussion can be more focused; therefore, presenters are encouraged to send materials to participants in advance. The notebook format of the Module allows the workshop leader to pull out posttests and simulations and reserve these for dissemination at the workshop. Pretests should be sent to participants ahead of time so a self-assessment can be completed.

### Use of Simulations

Because bias can occur at an individual and a system-wide level, simulations are designed to deal with both levels. For these simulations to be most effective, participants will want to examine bias in their own practices and in the policies of their school district. If a leader is presenting a workshop in his or her own school district, he or she may want to alter simulation exercises to deal specifically with the policies of the district. If participants represent several districts, the presenter may want to use the general simulations provided. Simulations may also be modified to focus on a particular minority population if this is more relevant to a particular area of the country.

For some simulations, participants may need to provide information on individual cases as well as information on district practices. When materials are sent out prior to the workshop, participants should be asked to bring whatever information may be necessary to participate in simulations, unless the workshop presenter chooses to provide it. This includes:

1. Information on the racial/cultural composition of the school population and the special education population in the school district.

2. A copy of the state rules and regulations for special education placement, including state and/or district definitions of mental retardation, emotional disturbance, and learning disability.

3. Information on a recent case in which a minority child was placed in a program for the mentally retarded; the emotionally disturbed.

4. Information on the district's language evaluation procedures and data on the language characteristics of the school population.

Other information may also be needed, and the workshop presenter should determine what else is necessary by carefully reading the simulations provided.

Finally, it is important that the presenter and participants be aware that the information in the Nonbiased Assessment Module will need to be supplemented and updated as court cases are settled, new legislation is introduced, and other information and techniques become available.

# NONBIASED ASSESSMENT: BASIC CONSIDERATION

Thomas Oakland
The University of Texas

# Table of Contents

Objectives ........ 1

Pretest ........ 1

A Diagnostic Intervention Process: The Components ........ 2

Biasing Factors ........ 2
  Parent Influences ........ 2
    Inadequately Informed Parents ........ 2
    Inconsistent Home and School Values ........ 3
    Unproductive Communication Systems ........ 3
  Child Characteristics ........ 3
    Language ........ 3
    Test Wiseness ........ 3
    Motivation and Anxiety ........ 3
    Cultural Differences ........ 3
    Expectations ........ 4
  Examiner Characteristics ........ 4
    Biased Attitudes ........ 4
    Deciding Who the Client Is ........ 4
    Competence ........ 4
  Diagnostic-Intervention Techniques ........ 4
    Referral ........ 4
    Screening ........ 6
    Assessment ........ 6
    Interpretation: Norms ........ 6
    Planning Inverventions ........ 6
    Executing Interventions ........ 6
    Follow-Up Evaluations ........ 6
  School System Policies and Practices ........ 6

Five Diagnostic-Intervention Models ........ 7
  Medical Model ........ 7
  Social System Model ........ 7
  Psychoeducational Process Model ........ 7
  Educational Task Analysis or Behavioral Model ........ 7
  Pluralistic Model ........ 8

Conclusion ........ 8

Posttest ........ 9

Simulations ........ 9

Key to Pretest ........ 15

Key to Posttest ........ 16

References ........ 16

Recommended Readings ........ 16

## Objectives for Nonbiased Assessment: Basic Considerations

The following constitute the objectives of this section of the module and are at the knowledge level:

- To know the various components of a good diagnostic-intervention program.
- To recognize the roles and influences of various persons in the success of diagnostic-intervention activities.
- To identify areas prior to, during, and following assessment where bias may occur.
- To understand the principal features of the five major conceptual models.
- To understand the functions of reliability, validity, and norms.

The following constitute objectives at the application level:

- To propose the development of a diagnostic-intervention program that considers the four major features.
- To characterize minority children for whom psychological tests may be valid and invalid.
- To describe the major features of a nonbiased assessment program for your school system.
- To identify the strengths, weaknesses, and applicability of the medical, sociological, psychological, educational task analysis and pluralistic assessment models.

With respect to your school system's current diagnostic-intervention policies, the goals include the following:

- Identifying possible sources of bias in each activity.
- Setting priorities for those areas in need of the greatest attention.
- Identifying areas over which you have some control.

## PRETEST

The following items are intended to assess your understanding of basic aspects of nonbiased assessment. The first 13 are short answer items, the remainder, True-False items. The answers are given in the key.

1. Identify two factors associated with a child which may be related to a teacher's
   (a) referring a child unnecessarily for assessment;
   (b) failure to refer a child for assessment when it is needed.
2. Identify four methods (other than IQ testing) that might provide useful information in a nonbiased assessment.
3. Define reliability and validity.
4. Define the term "norms" and list three norms against which a child might be evaluated other than national norms.
5. List two factors likely to increase the probability of successful interventions.
6. Cite two ways in which administrators may impede the development of special service programs.
7. List the major components of a well-designed diagnostic-intervention program.
8. Name two ways in which parent permissiveness may be detrimental to a child's educational success.
9. Define "test-wiseness" and list four of its components.
10. List three factors intrinsic to the child (unrelated to intelligence) which might affect the outcome of a standardized test.
11. Cite four aspects of a child's appearance which might bias a teacher's or psychologist's attitude toward the child.
12. Identify two factors which might prompt a psychologist to reduce allegiance to his/her client.

In front of the next 14 items place a T if you think the statement is true, F is you think the statement is false.

______ 1. No test is inherently biased unless its standardization sample did not include the specific client you are dealing with.

______ 2. Follow-up evaluations usually are one of the strongest features of a school's diagnostic-intervention activities.

______ 3. Though funding is important, it has little impact on the quality of nonbiased diagnostic-intervention techniques.

______ 4. PL 94-142 has had far-reaching effects on school policies generally but has not affected nonbiased assessment.

______ 5. In a well-designed diagnostic-intervention program, a child's medical history is important only when we have evidence that the child has been physically abused.

______ 6. A child's parents always can be counted on to do what's best for the child.

______ 7. Schools with a highly transient student population cannot provide a quality education.

______ 8. Child-screening methods vary considerably among school districts but the percentage of children found eligible for special services tends to be fairly stable.

______ 9. The WISC-R remains the best technique for an accurate, nonbiased assessment.

______ 10. Knowing the reliability of our data source is not necessary in nonbiased assessment when we are using criterion referenced assessment techniques.

______ 11. It is permissible to modify standardization procedures on an individual test if we have evidence that the child we are testing is multilingual.

______ 12. Cultural differences play a significant role in assessment except where English is the native language.

______ 13. It is acceptable to test a Spanish-speaking child in English when a bilingual education program is not available.

______ 14. Bilingualism may mask other problems such as speech difficulty.

## A Diagnostic-Intervention Process: The Components

The starting point for an appropriate nonbiased assessment is adoption of a good diagnostic-intervention model. Most school psychologists follow assessment formats which include what they believe to be the necessary components of the process; however their models may not be formally stated. Cromwell, Blashfield and Strauss (1975) provide a formal diagnostic intervention model which helps to conceptualize a realistic and viable approach for acquiring information that is important to understanding children and for using this information to develop successful interventions. It provides a good framework for our discussion of nonbiased assessment.

The discussion of this diagnostic-intervention model is limited here to its first four components labeled (A), (B), (C), and (D). The (A) and (B) components comprise the information-gathering and diagnostic side of the process. The first component (A) is the acquisition of *historical information* to help identify and understand important antecedent events in the child's life. The second component (B) involves collecting information that describes the child's current *characteristics* and those of the environment.

An adequate assessment model does not stop with information gathering, however. The third component (C) focuses on *interventions*. They can include any process over which we have control. Interventions are developed using information gathered from components (A) or (B) or both. The fourth component (D) estimates the success of the intervention, or the prognosis, given the information in (A) and (B).

Less complete diagnostic-intervention strategies often are used. Some strategies, for example, ABC, AC, and BC, are experimental and should be used cautiously; they base interventions on historical and/or current information but do not acknowledge the possible effects of the interventions. Progress must be monitored closely when one works with these strategies to insure that the interventions are working effectively.

ABD, AD, and BD diagnostic-intervention strategies are unsatisfactory because they do not provide for specific interventions. The prognosis for a 10-year old child with a history of academic problems, truancy, and current disruptive classroom behavior problems is made without describing the possible medical, social, psychological, or educational interventions which may be helpful.

Thus, the preferred model assesses historical (A) and current (B) conditions and provides for specific interventions (C) which are developed as an extension of the assessment information and which are likely to have a beneficial impact (D). The more our diagnostic-intervention model uses all four components (ABCD), the more likely it is to contribute to a quality nonbiased program.

Given this broad diagnostic-intervention model as a guide, more specific personal and administrative components of our work can be examined for possible sources of bias.

---

Task 1

*Define nonbiased assessment in your own words*

---

A quality *assessment* program can be judged on the basis of whether it provides the best available assessment models and techniques. A quality *nonbiased* assessment program can be judged on the basis of whether it provides a quality assessment program that eliminates, minimizes, or, at least, recognizes the presence of biasing conditions. Bias is apparent from predilections and procedures that prevent or obscure either (a) the full and accurate appraisal of conditions impeding a child's normal development or (b) the use of information to help maximize a child's development.

## Biasing Factors

There are many aspects of the diagnostic-intervention process and the people involved that can bias the assessment procedure. Characteristics of the parents, the child and the examiner, the adequacy of our techniques, and school system policies and procedures are discussed here as possible biasing factors.

### Parental Influences

Including parents as participants in the diagnostic-intervention process assumes that they will become active contributors. Most parents are eager to see their children get needed services, and their presence tends to have a beneficial influence. Unfortunately, others may be unwilling or unable to take an active role or may exert their influence in detrimental ways.

**Inadequately Informed Parents.** Some parents may not have the information necessary to help make appropriate decisions for their children. Some may know a Monday night T.V. schedule better than they know their children's school and after-school schedule. Others are caught up in the "Me" generation and lack a sense of dedication to their children. Some are unable to make objective, intelligent decisions regarding their children's welfare; are uncooperative, apprehensive, and afraid of the school; and may not have time or know how to help. These conditions may prevent or obscure the full and accurate appraisal of a child's needs.

The biasing role of parents who are *unwilling* to become adequately informed about their children's schooling may be difficult to overcome. Fortunately, however, in most cases parents are quite willing and will appreciate the school's efforts to provide them with information and, in turn, elicit from them ideas about their children's needs.

**Inconsistent Home and School Values.** Children have a right to education and society correctly expects this right to be exercised. Although the vast majority of parents highly value education, the actions of some are not consistent with allowing the school to be fully effective. For example, parents may encourage children to remain at home or to take a part-time job that escalates into full-time. Families may move frequently, disrupting children's educational, social, and emotional development. Schools that experience a 100% student turnover rate yearly are not able to provide quality educational programs nor to use acquired information to help maximize a child's development.

**Unproductive Communication Systems.** The communication system between the school and the child's family may be inadequate. Schools often use legalistic and educational terminology when they communicate with parents, perhaps in an attempt to comply with legal requirements. However, many parents often do not understand what is being said or implied and they feel confused and helpless. When asked to affirm school recommendations for their children's education, they may not comprehend the request fully. Such experiences may discourage them from attending other school meetings.

These problems are compounded when a parent's English proficiency is limited. Some parents who are not native English speakers acquire enough English to show survival skills yet they may not grasp the nuances and complexities of educational terminology.

Many minority parents of today were yesterday's children who experienced the inequities and inadequacies of educational programs which PL 94-142 has sought to remedy. As children, some of these parents were placed in low-ability classrooms or were labeled inferior or deviant because of cultural and language differences. These early experiences may color their current attitudes toward schooling. For such parents school personnel must exercise skillful and empathic efforts to help them to perceive their children's educational experiences in a different light.

These and other parental influences may impede a nondiscriminatory program.

---

*Task 2*

*Suggest ways that a school system may acquire needed family information and support, given obstructive parents.*

---

**Child Characteristics**

**Language.** Children's ability to understand and communicate in English is very important. School success depends in part on being able to understand, speak, read, and write English. Thus it is important to know children's English proficiency in order to judge whether their language skills are sufficiently developed to enable them to perform adequately on tests. Conventional tests that require a high level of English proficiency cannot be used with children whose English language skills are poorly developed. Paucity of language may signal a general language deficiency or a language difference owing to either an exposure to nonstandard dialects or exclusive knowledge of a language other than English.

The tests used in schools generally are not intended to assess language skills directly but to use language as a means of assessing intelligence, achievement, personality, and other characteristics. Conventional assessment techniques often must be altered when we work with children with language differences so that we can eliminate biases arising from language and develop valid profiles of their academic, intellectual, personality, and social characteristics. (See section of module by Matluck & Mace-Matluck for a detailed discussion of language characteristics and nonbiased assessment.)

**Test Wiseness.** We administer tests to children with the presumption that they already have acquired certain requisite abilities and attitudes. For example, we assume that they understand directions (which may include concepts such as right, left, up, down, same, and different), that they consider all possible responses before choosing the correct one, that they work on one item at a time and are not distracted by other items, and that they are involved and attentive during the entire test. These and other abilities (Oakland, 1972) constitute basic test-taking skills. A lack of test-wiseness contributes to bias. Thus, we must insure that children have prerequisite test-taking skills.

**Motivation and Anxiety.** Adequate test performance requires that children be properly motivated (Havighurst, 1970). Results from aptitude and achievement tests are valid only when children are performing at their very best. Too often children randomly select answers on a multiple choice test, fail to cooperate, and show their lack of motivation in other ways. Other children may be extremely anxious and unable to concentrate and attend to the test. A nonbiased assessment program must take into account the attitudinal characteristics of children to insure that they are properly motivated.

**Cultural Differences.** Children often come from restricted or different physical and cultural settings where the opportunities for growth and development differ significantly from those available to most children (Cole & Brunner, 1971; Newland, 1973). These differences may be seen in child-rearing practices, expectations and aspirations, language experiences, informal and formal learning experiences, and other elements of acculturation patterns. The acculturation patterns of minority group children and children from lower socio-economic homes *may* be significantly different from the patterns of children who are included in a test's standardization sample. Confidence in using a test decreases when a child's acculturation patterns differ significantly from the patterns which are normally provided for other children.

Youngsters do not significantly differ from each other solely by virtue of minority or lower socio-economic group affiliation. The decision on whether a child's acculturation patterns are similar to those of other children who are included in a test's standardization sample can be made for *each child individually* only after thorough knowledge is obtained of each child's background and of the test's standardization sample. (See section by Henderson in this module for a discussion of sociocutural determinants of behavior.)

---

*Task 3*

*Briefly describe two black children you may have worked with, one with whom the WISC-R can be used and one with whom the WISC-R cannot be used.*

---

**Expectations.** A person's behavior tends to move toward the expectations for him/her which other people hold. When we expect people to be well behaved and we communicate these expectations to them, the prevalence of good behavior increases. Children tend to adopt and accept the expectations which their peers, family, and teachers communicate to them. Knowing those expectations for a child enables us to appraise his or her future more accurately and alter our interpretations of the assessment data. A child who expects to fail on tests will underperform on them. Thus, low expectations can add to our bias and exacerbate other problems.

---

*Task 4*

*List those examiner characteristics that may cause significant problems in working with minority group children.*

---

**Examiner Characteristics**

The assessment specialist also plays a central role in designing and carrying out a nonbiased assessment program. Three particular areas are of great importance.

**Biased Attitudes.** We know that adults' attitudes can affect their behaviors toward children. Psychologists who feel attachment for and are concerned with students behave quite differently from those who feel rejection or indifference toward them. Moreover, psychologists' attitudes can be affected by children's characteristics. School personnel generally tend to favor bright, achieving, linguistically competent, academically motivated, compliant, conforming students. However, many children who are referred for assessment exhibit quite different characteristics. Furthermore, some persons have strong and fixed opinions of persons of identifiable racial-ethnic and social class groups. These prejudices act as blocks to prevent people from understanding the individual characteristics of one such person. School psychologists are not immune to these prejudices. They, too, may have biases. Thus, it is important to assess the extent to which bias may discolor and alter both the information we acquire and the interpretations we make of this information when we work with persons of different racial-ethnic and socio-economic groups.

**Deciding Who the Client Is.** The American Psychological Association (1972) emphasized the belief in the dignity and worth of the individual, a commitment to freedom of inquiry and communications, and a concern for the best interests of clients, colleagues, and society in general. Psychologists are strongly encouraged to respect the integrity and protect the welfare of the persons with whom they work. When a conflict arises among professional workers, psychologists should be concerned more with the welfare of their clients (e.g., children) than with the interests of their professional group.

This principle, however, is not always adhered to. Some examiners are more concerned with job security, friendships, and serving the school system that employs them. Persons working within this frame of mind will not fully investigate all school-related factors that may attenuate a child's performance. To find a deficit within the child and to fault the child's home and neighborhood often is easier than to identify important teacher- and other school-related variables that impede the child's development. In a nonbiased assessment program, the examiner has an open mind and investigates both school- and home-related factors that may be hampering a child's development.

**Competence.** Examiners tend to be highly trained, competent, and dedicated. However, some know assessment superficially and mechanically, are poorly prepared in psychoeducational, child clinical, and behavioral assessment (see section of this module by Tombari for a discussion of behavioral assessments for emotionally disturbed students), and generally have not kept up with the advancements made in the field of appraisal. Other psychologists consider their full-time job to begin after school and they devote more effort to developing their private practices than their school-related activities.

Many minority group parents and those with low incomes depend upon the public schools to provide quality educational and psychological services. These parents do not have the financial means to purchase such services privately. Thus, the standards governing the provision of educational and psychological services in public schools must remain as high as those for the private sector. Nonbiased assessment is based on employing people who are competent and on providing high-quality services to children and their families through the public schools.

### Diagnostic-Intervention Techniques

Various sources of bias may be woven into our diagnostic-intervention process. The detection of some forms of bias is increased when we first conceptualize the information-gathering phase (i.e., C and D) as composed of three components (planning interventions, executing interventions, and conducting follow-up evaluations and interventions). Each component is susceptible to certain influences that limit its effectiveness.

---

*Task 5*

*For each activity list one or more sources of bias which may have prevented you from working effectively with a minority group child.*

*Referral:*
*Screening:*
*Assessment:*
*Interpreting Information:*
*Planning Intervention:*
*Executing Intervention:*
*Follow-up:*

---

**Referral.** The diagnostic-intervention process begins when a teacher refers a child for special attention. The nature of the behaviors that actually stimulated the referral may be influenced by bias. A teacher may refer a child for having an academic problem when, in fact, the child may not have a severe academic problem but, instead, exhibits other behaviors which the teacher finds disturbing. Some teachers have lower expectations for boys; for residents of tenements or mobile homes; for children who come from lower class homes, attend unconventional churches, dress poorly, have one-parent families, speak a foreign language, arrived in town recently, or have other distinguishable characteristics. A child's skin color or last name may stimulate other deep-seated prejudices. Prejudices may result in the identification of qualities that are not present and encouraging the development of latent qualities. Important characteristics are often overlooked, for example, the speech problems of children with bilingual backgrounds or of those who speak nonstandard dialects. Health problems that teachers readily detect among middle-class children may go unrecognized in lower class children.

A teacher who wants a child removed from the classroom, refuses to try different educational and behavioral strategies, distorts information, is obstructive and uncooperative, or thinks that the child's environment is so deleterious that nothing will be beneficial, also stands in the way of developing suitable nonbiased programs.

**Screening.** Administrative procedures for processing referrals differ from district to district. In some districts children's names are put on a waiting list for screening and evaluation in turn by the psychologists and other diagnostic specialists. In other districts names are referred to within-school committees that are responsible for screening referrals, collecting existing data, obtaining additional information, and identifying the resources in the school that can be used to meet children's needs. Clearly, the latter approach is preferable, but this process too must be carefully monitored for biasing factors.

In some districts over 90 per cent of children referred for special services are found to be eligible. Imbalances in the number of minority group children placed in special education may begin at this point. Thus, an examination of referral and screening procedures may indicate whether they are significant sources of bias.

**Assessment.** Nonbiased assessment occurs only when the best available methods and techniques are used. During the last 10 years discussions of nonbiased assessment have been dominated by the WISC-R. Consequently, we may have lost sight of the full complement of methods that are useful to the gathering of information. Some good child-study methods include the following:

- Obtaining behavioral data.
- Employing structured and unstructured interviews with children, teachers, family members, and other significant persons.
- Using existing sources of data (e.g., school and medical records).
- Requesting additional data.
- Obtaining information directly from children through observations at home and school.
- Assessing the child through informal or formal techniques, including criterion- and norm-referenced measures.

Other methods are useful for understanding the effects that home and school factors exert on a child's behavior. Each, together with other techniques, has a place in a nonbiased assessment program. The selection and use of assessment methods depends in part on the quality of information each provides. Quality information requires that our source be reliable and valid.

It is difficult to discuss reliability and validity separately. Both are essential characteristics of good measurement instruments and, as such, are particularly crucial if assessment is to be nonbiased. Reliability refers to the stability and consistency of the measurement. If the instrument or technique used to measure some dimension of a person is not reliable, it cannot be assumed that the person being assessed would earn the same score or rating if measured again. A measure which is not reliable cannot be valid.

The validity of a test indicates how well it measures what it is designed to measure and how useful it is in predicting other behaviors or outcomes. Statements or predictions based on invalid tests are meaningless and foster inaccurate decisions. It is particularly important, however, to be aware that although a test may be valid for one purpose or group it may be invalid for another.

Tests which lack validity have increased potential for biasing assessments. Predictive or criterion-related validity and content validity are particularly important in nonbiased testing and reporting of results. Using a test which lacks criterion-related validity for a particular cultural group may lead to inaccurate placement decisions or predictions about students from that group. Bias may also occur if items (content) of the test do not represent comparable tasks to members of different cultural groups.

Test users should not, however, rely on their personal judgment to determine whether bias is present in particular items or tests. It is their responsibility to examine test manuals and select tests which are reliable and valid for their purposes. One advantage of norm-referenced measures is that they usually report reliability and validity data. Statistics such as the standard error of measurement, the standard error of estimate and the standard error of difference are useful in estimating the reliability and validity of a particular child's score. The standard error of measurement is also important in reporting and interpreting scores in a nonbiased manner.

**Interpreting Information: Norms.** Against what standards do we judge a child? We have at least three broad criteria: perfection, potential, and par.

Some teachers and parents expect children's work and behavior to be perfect and error free. However unrealistic such a standard may be, it allows us to objectively evaluate the children's performances. If the children behave perfectly, they are meeting the criterion; if they are not meeting the criterion, perhaps more realistic standards are needed.

Standards frequently are established by estimating what children are capable of doing on the basis of their potentials. Frequently, IQ tests are given to help to estimate children's academic potential. Boards of education throughout the U.S. have faced angry parents who expressed displeasure that their children's academic potentials were not acknowledged and who opposed placement in lower ability groups.

A third standard evaluates one child against other children, that is, compares a child with a peer group to see if he or she is up to par. Questions such as, "Is it normal for Jose to act this way?" or "Why is Josephine still reading at the primer level while the others are in their third ad fourth readers?" are asked with certain norms in mind. Most of the standardized tests we use in assessing children use this kind of comparison to make interpretations about performance.

Evaluating a child against standards, based on the performance of other children requires that we know the characteristics of the other children and whether our child is a member of the same population. Standards can be stated in terms of norms for class, campus, city, state, regional, national, or international populations.

One assumption underlying norm-referenced assessment is that the child being tested could have been included in the test's standardization sample (i.e., the children whose scores are used to make up the test norms). For many tests the norms are large, heterogeneous, and well selected to adequately reflect the full range of children's characteristics; for other tests, the standardization sample is restricted and narrow.

No test is inherently biased. Bias enters when someone uses a test with an inappropriate standardization sample for purposes that cannot be supported by the validity data. '

In order to use a norm-referenced test with minority group children, the test's standardization sample should include children who are similar in age, grade, gender, socio-economic status, race or ethnicity, and cultural experiences. Knowing the standards against which a child's performance is being compared is prerequisite to a sound nonbiased assessment program.

The development of a successful nonbiased assessment program requires that work also be directed toward developing good interventions. What happens to a child following assessment has a direct influence on the success of the assessment program. Three principal components must be considered for a proper program following assessment:

planning, executing, and evaluating interventions. Each is considered in turn.

**Planning Interventions.** The primary goal of assessment is to amass accurate data which can be used as the basis of beneficial decisions and interventions. The planning of interventions is likely to be effective under the following conditions:

- Information covers the various characteristics (e.g., home, school and child), is reliable, and describes precise behaviors.
- Recommendations are made by the persons who provide information and are selected because they are known to be effective.
- Interventions include the possibility of modifying both the child and the school environment.
- Responsibilities for the interventions are assumed willingly by people with appropriate authority.

Large numbers of minority group children have been routinely placed in lower ability groups, EMR classes, and other administrative structures which are ineffective and inferior to regular education programs (Oakland & Laosa, 1977). For example, in the early 1970s, blacks constituted 9% of California's population and 26% of the educable mentally retarded population. Given the strong and pervasive notion that EMR classes are educational dead ends and ineffective, many persons believe that the disproportionate assignment of minority group children to these classes provides *prima facie* evidence for discriminatory practices (see section in this model by Bersoff for discussion of legal points). Care must be taken to insure that judgments of the most suitable placements and interventions also take into account the proportionate assignment by race to special education categories.

**Executing Interventions.** All the activities discussed so far lead to this activity. Offering interventions which are designed to facilitate children's development is our most important activity in a nonbiased assessment program. Yet, in all too many situations, specialized interventions do not occur. Although a child may be relabeled and placed in a different class, important and more effective curricular and behavioral strategies do not routinely follow.

Various factors contribute to this ineffectiveness (see preceding discussion of parent, child, examiner, and administrative characteristics). Strong sources of bias may also exist at a systems level. School system policies and practices that contribute to bias are discussed in a following subsection.

---

*Task 6*

*What policies and practices within your school system seem to inhibit or impede the carrying out of specialized intervention programs for children?*

---

**Follow-up Evaluations and Interventions.** Here we meet what is, perhaps, the weakest feature in diagnostic-intervention activities. Follow-up evaluations, when they occur, often are performed in a perfunctory manner with little attention to acquiring the data that are useful for educational or behavioral programming.

This phase corresponds to the prognosis (D) component in our diagnostic-intervention model. Reassessments provide the additional information that is needed to evaluate the success of the various interventions. Three questions are asked at this point:

- How much progress, if any, has been made?
- What factors have exerted a positive or negative impact?
- How can the interventions be altered to improve their effectiveness?

**School System Policies and Practices**

Probably the most important influence in a nonbiased assessment program is the extent of the school system's financial and professional resources and the degree to which it is willing to commit them.

School district policies often promote the development of bureaucratic networks which are insensitive to the individual needs and characteristics of children, their families, and teachers. During the last 20 years the federal government has actively shaped school policies and programs through legislation and judicial decisions. Their influence at times is beneficial but formulating school policies to comply only with federal law does not encourage us to provide quality professional services to children. School districts that identify many children for special education programs in order to insure a full share of state and federal funds illustrate one common aspect of improper compliance.

Many school districts are financially troubled, constantly face the threat of strikes and other disruptions in their programs, and have a high turnover rate and low morale among professional staff and children. These problems hinder attempts to develop programs which are suitable for children from diverse racial-ethnic, cultural, and social groups.

The influence of Public Law 94-142 has been significant and far-reaching on education generally and on nonbiased assessment specifically. (See section in this module by Bersoff for a discussion of legal issues associated with nondiscriminatory assessment.) One effect has been to encourage school psychologists to assess more children directly and not to provide consultation services to teachers, principals, and parents. In some communities the number of children being referred for appraisal has dramatically increased because of a decrease in consultation activities. Consultation permits school psychologists to work directly with teachers and parents to arrive at viable solutions for educational, social, and psychological problems without doing complete appraisals of children, thus permitting greater numbers of children to be seen.

Some legal requirements seem beyond the capabilities of school districts. In New York and Chicago children speak more than 200 languages and dialects. Although schools are responsible for assessing children in their native languages, this obligation probably cannot be met completely.

Administrators impede the development of good programs when they do not provide sufficient support for special services; are uncomfortable working with children with special handicaps, minority groups, or persons from low socio-economic status homes; evidence dictatorial leadership styles rather than facilitative problem-solving styles; and draw rigid boundaries between school and neighborhood, thus inhibiting teachers and parents from forming important and mutually supportive relationships for the welfare of children.

Clearly, problems in assessing minority group children arise from more than one source: parent, child, examiner, assessment characteristics, and general system influences can be named. Thus, the design for a nonbiased assessment program must consider the influences of all these factors.

---

*Task 7*

*Thinking about your school system and its nondiscriminatory assessment program, set priorities for those areas that need the greatest attention. Put a 1 next to the area needing* *the most immediate attention, a 2 next to the second most critical area, etc.*

_____ *parent*
_____ *child*
_____ *examiner*
_____ *referral*
_____ *screening*
_____ *assessment*
_____ *interpreting information*
_____ *planning interventions*
_____ *executing interventions*
_____ *follow-up evaluations*
_____ *system policies and practices*

---

Knowing the various sources of bias that may scuttle attempts to offer a quality nonbiased assessment program may leave us feeling overwhelmed and powerless.

---

*Task 8*

*Return to the preceding list and identify those areas over which you have some control. Put an A next to the area over which you have most control, a B next to the second area, etc.*

---

By identifying the areas over which we have some control and authority, we more clearly delineate important areas in which we can influence events. This knowledge can be particularly important when our As and Bs coincide with the ones and twos in the preceding list.

## Five Diagnostic-Intervention Models

Historically, we have come to accept a few axioms regarding diagnostic-intervention programs. One is that important differences exist both *between* persons and *within* persons. Describing and allowing for these differences presumably helps us to describe a person more accurately, to develop plans based upon between and within individual differences, and to promote the person's development more effectively. Moreover, we believe an extensive and complete examination is more helpful than one in which a few characteristics are examined narrowly.

Behaviors often are viewed according to four different models: medical, sociological, psychological, and educational (task analysis or behavioral). Each model has techniques that permit us to examine children's inter- and intra-personal differences (Mercer & Ysseldyke, 1977). More recently, in response to the challenge on assessing minority group children, a fifth model has been offered: pluralistic assessment (Mercer & Lewis, 1978). Each model differs in such characteristics as how normal and abnormal behaviors are defined and assessed; assumptions about behaviors; how data are collected, interpreted, and used; and the extent to which social class and racial-ethnic differences bias the data.

School psychologists are likely to be involved in a full range of diagnostic-intervention activities at the same time that they focus on assessment and related activities, including referring, screening, and interpreting information.

Quality assessment is the cornerstone of a nonbiased assessment program. By knowing the assets and limitations of each model, we can use assessment techniques more appropriately and confidently. Each of the five models, when used appropriately, contributes to a diagnostic-intervention program that eliminates, minimizes, or at least recognizes the presence of biasing conditions.

### Medical Model

The medical model provides for ABCD and BCD relations. Abnormal characteristics are detected by the presence of biological symptoms which, presumably, are caused by biological conditions. Once the deficit is detected, treatments are prescribed to improve the biological organism. Measures, such as for vision, hearing, and heart rate, are designed to assess biological symptoms and are validated through comparisons with other biological data. Medical measures provide dichotomous data (e.g., the symptom is either present or absent) which are interpreted as information about a person's biological make up. Racial and cultural discrimination is minimal provided test interpretations and treatment are confined to the biological organism.

### Social System Model

This model most frequently employs BCD or BD constructs and emphasizes learning and the display of socially approved behaviors. Abnormal behaviors are those that violate social expectations for specific roles whereas normal behaviors conform to social expectations. Thus, no one definition or standard describes what is normal or abnormal. Definitions for normal behavior depend on knowing the expectations of various people for the roles others occupy in specific situations. For example, whether spitting is an appropriate activity depends on the setting and on the role of the spitter; a child who spits at school may be punished, but one who spits in the dentist's office is reinforced; and in the dentist's office the client is encouraged to spit but not the dentist and receptionist. Thus, social systems define appropriate behavior according to the person's role and the place in which it occurs.

Social measures of such characteristics as adaptive behaviors and social development are used to assess competence in social roles and are validated by correlating them with the judgments of other persons. The measures are used to discover deficits and assets. Scores are multivariate and normally distributed (i.e., distributed continuously, rather than dichotomized), and interpreted by describing and evaluating a person's role in a specific social system. Racial and cultural bias and discrimination are minimized or avoided by describing and specifying the person's role and the social system when interpreting scores and suggesting interventions.

### Psychoeducational Process Model.

This model relies heavily on BC(D) relationships and emphasizes assessment and training in the use of specific processes or abilities, often through a compensatory or remedial program. Assessment is directed toward identifying deficient processes or abilities that, presumably, are related to and the cause of a child's academic difficulty. Cognitive, perceptual, psycholinguistic, and psychomotor measures provide multivariate, normally distributed scores that are interpreted against standards established by the test's standardization sample. A high test score suggests high abilities and eliminates the process or ability as a cause of the problem; high scores also may be used to identify talented and gifted children. A low test score identifies the ability as a probable cause of the problem. Racial and cultural bias and discrimination are minimized when the characteristics of a test's standardization sample match those of the children being tested. Differences between the acculturation patterns of a child being assessed and those of the children in the standardization sample increase the likelihood of test bias. As with all tests, psychoeducational measures should have sufficient items to sample adequately the domain being assessed. Suggestions for ineffective interventions (e.g., "to improve general perceptual, intellectual, or linguistic processes") may be common in this model and constitute a strong source of bias.

### Educational Task Analysis or Behavioral Model.

BC(D) type constructs are used in a test-teach-test activity. A child's behavior is the result of the interaction between a set of enabling behaviors and the task requirements. Comparisons that involve definitions of normal and abnormal are avoided. A child's behaviors are evaluated in terms of an apparent ability or inability to perform certain tasks.

The only cause presumed for an inability is the absence of lower level enabling behaviors. That is, the model assumes that there are skill hierarchies and the development of higher level and more complex skills is dependent on the adequate development of lower level skills.

To be appropriate, measures must be based on the skills, abilities, and behaviors a particular environment endeavors to foster. They must identify what a child does and does not do and they should parallel classroom resources that encourage children to learn the necessary enabling behaviors. Criterion-referenced measures and behavioral observation techniques frequently are used; they provide dichotomous data. Each child's score is used to evaluate his/her development alone; it is not compared with other children's scores. The skills (and thus the tests) relate to the school's curriculum and subject matter. Racial and cultural bias and discrimination are absent when each score is interpreted in terms of a child's development on a well-sequenced curriculum that provides for the continuous achievement of enabling behaviors and skills.

**Pluralistic Model.**

The least well-developed of the five models, the Pluralistic Model, emphasizes BD constructs and uses various techniques to discover the talent and potential which are masked by culturally biased and inappropriate measures. Learning potential is assumed to be similar among all racial-ethnic and sociocultural groups. Thus, scores from tests that show racial-ethnic and sociocultural differences are assumed to be biased and in general need of adjustment.

Three general procedures can be used with this model to eliminate bias. One procedure is to develop culture-specific tests that measure a specific aspect of culture and are standardized on a well-defined and usually narrow sample. A second procedure attempts to estimate a child's learning-to-learn ability or the ability to profit from experience. For this procedure, a child is pretested, taught some new and relevant skills, and posttested. The amount of gain the child shows between the two tests is the index of potential. The third approach (Mercer & Lewis, 1978) uses multiple test norms for children from various sociocultural, socioeconomic, racial-ethnic, and geographic groups.

Pluralistic norms are thought to be needed on an instrument if the mean scores differ for children from different racial-ethnic groups, if children from these racial-ethnic groups also differ in social, economic, or cultural characteristics, and if the social, economic, or cultural characteristics correlate with test performance. Tests that do not exhibit these three conditions will evidence little bias when they are used without pluralistic norms.

A child's score is interpreted using only the relevant test norms. For example, Mercer and Lewis's (1978) estimated learning potential adjusts a child's WISC-R score upward to reflect his/her performance relative to other children from similar sociocultural and racial-ethnic groups. The test data are multivariate and distributed somewhat normally. A child's score is compared to children of similar sociocultural and racial-ethnic characteristics and may be below average, average, or above average.

Measures used in the pluralistic model often were developed originally for use in the social systems and psychoeducational models. Following their reinterpretation in light of pluralistic procedures, the measures often are reapplied in their original models, but now with adjusted scores that are derived from using a more tightly specified norm group. Racial-ethnic and sociocultural bias are eliminated by precisely specifying a norm group of which a child is a member, by interpreting scores within that norm group, and by eliminating inferences about how a child would perform relative to other norm groups. Few guidelines exist on how pluralistic data can be used to develop interventions. Thus, the current models in this model have limited value for helping us to develop a nonbiased diagnostic-intervention program.

## Conclusion

A nonbiased assessment program provides information that is useful for developing and carrying out interventions in ways that eliminate, minimize, or recognize the presence of biasing conditions. Various conditions can bias our predilections and procedures and so prevent or obscure either the full and accurate appraisal of conditions that impede a child's normal development or the use of information to help maximize a child's development.

In a more proactive way, we can characterize a nonbiased program as one that helps people to make wise and informed decisions; seeks to improve the development and status of all persons; and identifies and fosters the development of various talents, abilities, and characteristics in people.

## POSTTEST

1. List and describe the first four components of the diagnostic-intervention model proposed by Cromwell, Blashfield, and Strauss.

2. Name five possible sources of bias in diagnostic-intervention activities.

3. Discuss two reasons that may account for a child's low language functioning.

4. Discuss some methods that might be effective in helping a child to acquire basic test-taking skills.

5. Discuss three reasons why the follow-up evaluations and interventions ("D" component in the diagnostic-intervention model discussed in the text) may be the weakest feature of diagnostic-intervention activities.

6. Describe the essential features of the medical, sociological, psychological, educational (task analysis or behavioral), and pluralistic assessment models.

7. What kind of ineffective interventions are likely to be recommended by the psychoeducational process model?

8. Describe the assumptions that underlie norm-referenced testing with psychological tests. How do these assumptions relate to nonbiased assessment?

## SIMULATION 1

This simulation requires you to consider information from a case involving a minority child with whom you worked recently. Whenever possible, obtain a copy of the case folder. Be sure to delete all names and other identifying materials in order to maintain the confidentiality of this information. You will not be asked to share any case information with other workshop participants.

Refresh your memory of the specific details of this case by reading the case folder. After reading the case, and knowing the standard operating procedures in your school system, complete the following questions. Circle those questions which you feel constitute particular weaknesses with this case or are typical of your system. Following this activity you and other workshop participants are encouraged to describe ways to eliminate impediments to effective nonbiased programs.

### Referral[1]

1. Were the parents/guardians aware that a referral had been made for their child, and by whom?
2. Is this child's presenting problem clearly and precisely stated on the referral?
   a. Does the referral include descriptive samples of behavior rather than opinions of the referring agent?
   b. Is there supportive documentation of the problem?
3. Is the referral legitimate?
   a. Does the referring agent have a history of overreferral of children from certain cultural or racial-ethnic groups?
   b. Could irrelevant personal characteristics (e.g., sex or attractiveness) of the child have influenced the referral decision?
   c. Could the referring agent have misinterpreted the child's actions or expression because of a lack of understanding of cultural differences between self and child?
4. Did the assessment team provide the referring agent with interim recommendations that may eliminate the need for a comprehensive evaluation?
   a. Is it possible that the curriculum being used assumed that this child has developed readiness skills at home which, in reality, he hasn't had the opportunity to develop? If so, did the team assist the teacher in planning a program to give this child the opportunity to develop readiness skills?
   b. Did the team provide information on the child's cultural background for the referring agent so that there are fewer misunderstandings between the referring agent and this child and perhaps other children of similar cultural background?
5. If necessary, were the child's parents/guardians informed of the referral in their primary language?
   a. Were the reasons for the referral explained?
   b. Was there a discussion with the parents on what the next activties may be? e.g.,
      —professional evaluations
      —use of collected data
      —design of an individualized education plan, if necessary
   c. Were due process procedures discussed with the parents?
   d. Is parental permission for the evaluations documented?
   e. Have the parents been asked to actively participate in all phases of the assessment process?
   f. Have the parents been informed of their right to examine all relevant records relating to the identification,

evaluation, and educational plan of their child?

---

[1] The following questions are adapted from "A Position Statement on Nonbiased Assessment of Culturally Different Children" prepared by the Region 9 Task Group on Nonbiased Assessment, sponsored by the Northeast Regional Center, Hightstown, N.J.

## Meeting the Child

1. What special conditions about this child were considered?
   a. What is the child's primary home language?
   b. What do we know about the child's home environmental factors? e.g.
      —familial relationships/placement
      —social and cultural customs
   c. Do I understand this child's culture and language so that I can evoke a level of performance that accurately indicates the child's underlying competencies?
   d. Is this child impeded by a handicap other than the referral problem that may result in his not understanding what is discussed?

2. What special conditions about me do I need to consider?
   a. How did I feel about this child?
   b. Are my values different from this child's?
   c. Will my attitude unfairly affect this child's performance?
   d. Could I evaluate this child fairly and without prejudice?
   e. If not, would I refer him to another assessor if one is available?

3. Did I examine closely all the available existing information and seek additional information concerning this child?
   a. Has the child's academic performance been consistent from year to year?
   b. Is there evidence in this child's record that his performance was negatively or positively affected by his classroom placement or teacher?
   c. Are his past test scores consistent with his past class performance?
   d. Am I familiar with past test instruments used to evaluate this child and how well can I rely on his prior test scores?
   e. Has this child been observed in other environments (individual, large group, small group, play, home)?
   f. Were illegitimate assumptions made about this child (e.g., is it assumed he speaks and reads Spanish simply because he is Puerto Rican)?
   g. Was information sought on non-school related variables that may have affected this child's school performance? e.g.,
      —health factors (adequate sleep, food)
      —peer group pressures

4. Did the child understand why he was being assessed?
   a. Did I try to explain it at his level of understanding?
   b. Was the child given an opportunity to freely express his perceptions of "the problem"?
   c. Did I or someone discuss with the child what next activities may be involved?

## Selection of Approach for Assessment

1. Was the best assessment approach used for this child?
   a. Considering the reasons for referral, did the techniques use behavioral observations, interviews, informal techniques, or standardized techniques, or a combination of the above?
   b. Was adaptive behavior assessed thoroughly?
   c. Were the approaches consistent with the child's receptive and expressive abilities?
   d. Was there an overdependence on one technique, overlooking others that may have been more appropriate?
   e. Was there a balance between formal and informal techniques?

2. If standardized instruments were used, what ramifications were considered?
   a. Was this child tested simply because we always used these tests in the assessment procedure?
   b. Was a particular test readministered simply because it is part of THE BATTERY?
   c. Was a test administered because of a directive from the Administration?
   d. Do the instruments' standardization samples include persons from the child's racial-ethnic and cultural groups?
   e. Are subgroup scores reported in the manual?
   f. Were there large numbers of children from this child's racial-ethnic and cultural group included in the test standardization sample?
   g. Do the instruments selected assume a universal set of experiences for all children?
   h. Do the instruments selected contain illustrations that are misleading and/or outdated?
   i. Do the instruments selected employ vocabulary that is colloquial, regional, and/or archaic?
   j. Is the theoretical basis of the instruments understood?
   k. Will this instrument easily assist in delineating a recommended course of action to benefit this child?
   l. Has current literature regarding this instrument been reviewed?
   m. Has there been a review of current research related to potential cultural influences on test results?

### Test Administration

1. Were there factors (attitude, physical conditions) that supported the need to reschedule this child for evaluation at another time?
2. Could the physical environment of the test setting adversely affect this child's performance?
   - —room temperature
   - —noise
   - —poor lighting
   - —furnishings inappropriate for child's size
3. Am I familiar with the test manual and did I follow its directions?
4. Was this child given clear directions?
   - a. If his native language is not English, was he instructed in his language?
   - b. Is there assurance that this child understood the directions?
5. Are entire responses to test items accurately recorded, even though the child's answers may be incorrect, so that they can be considered later when interpreting his test scores?
6. Was rapport established and maintained with this child throughout the evaluation session?

### Scoring and Interpretation

1. Has each item missed by this child been examined rather than merely his total score looked at?
   - a. Is there a pattern to the types of items this child missed?
   - b. Are the items missed free of cultural bias?
   - c. If all items thought to be culturally biased were omitted, would this child have performed significantly better?
2. Were various factors considered in the interpretation of this child's scores?
   - a. The effect the child's attitude and/or physical condition had on his performance?
   - b. Do the interpretations of this child's performance include observations?
   - c. Are scores reported and interpreted within a range rather than as one number?
3. What confidence can we have in this child's test scores?
   - a. Are test scores the most important aspect of this child's evaluation?
   - b. Do the test scores outweigh professional judgment about this child?

### Consultation with Team Members and Others

1. Is there evidence that a committee functioned as a multidisciplinary team on behalf of this child?
   - a. Did all professionals on the team share their findings on this child?
   - b. Are team members' evaluations in conflict?
   - c. Can each member admit his discipline's limitations and seek assistance from other team members?
   - d. Do the professionals willingly share their competencies and knowledge with other team members for the benefit of this child?
   - e. Has the team arrived at its conclusions as a result of team consensus or was its decision influenced by the personality and/or power of an individual team member?
2. Is the multidisciplinary team aware of its limitations?
   - a. Are they aware of community resource personnel and agencies that might assist them in developing an educational plan for this child? Were such resources used before, during, and after the evaluation?
   - b. Did the team feel comfortable in including this child's parents in their discussions?
   - c. Did the team engage in discussions that did not include the child's parents?

### Assessment Report

1. Is the report clearly written and free of jargon so that it can be easily understood by the child's parents and teachers?
2. Does the report answer the questions asked in the referral?
3. Are the recommendations realistic and practical for the child, school, teacher and parents?
4. Have alternative recommendations been provided?
5. Does the report describe any problems that were encountered and the effects of such during the assessment process?
6. Do the recommendations provide for possible modifications of the child, teacher, curriculum, and parents?

### Individual Educational Plan

1. Is the child made to fit into an established program or is an individualized educational plan that is appropriate for this child being developed?
   - a. Are this child's strengths and weaknesses identified?
   - b. Are long-range goals and immediate objectives specified for this child?
   - c. Are the team members willing to assist the teacher in carrying out this child's educational plan?
   - d. Does the report state when and how this child's progress will be evaluated and by whom?

### Follow Up

1. What are our responsibilities after we have written this child's educational plan?

a. Have the findings and recommendations been discussed with this child's parents and their due process rights explained? Have the parents been given a written copy of this child's educational plan?
b. Have the persons working with this child met to discuss the educational plan and to assist one another in implementing its recommendations?
c. Have the findings and recommendations been discussed with this child at his level of understanding?
d. Can I help the persons working directly with the child to become more familiar with this child's social and cultural background?
e. Have the parents granted permission for releasing any confidential materials to other agencies and professionals?
f. Will this child's educational plan be reviewed periodically in regard to his actual progress so that necessary changes can be made?

## Some Final Thoughts

1. To what extent do I believe in the right to an appropriate education for all children?
2. To what extent would I be comfortable if MY child had been involved in THIS assessment process?
3. To what extent am I willing to actively participate in inservice activities that will lead to the further development of my personal and professional growth?

## SIMULATION 2[1]

Pedro was referred to a child study team of Pecan Elementary School in January. The team members consisted of the school counselor, the three first-grade teachers who team teach, a school psychologist, and a special education supervisor. Pedro's teachers provided the following information.

Pedro (CA=7) is enrolled in the first grade. All the first-grade team teachers are concerned about Pedro's lack of academic progress. He has acquired very few readiness skills, and his progress is slower than most of his classmates. He cannot name letters, does not know their sounds, cannot count, and barely is able to write his name. Pedro also exhibits articulation difficulties, and his speech is difficult to understand. He comes from a bilingual background; both English and Spanish are spoken at home. The bilingual aide has indicated that Pedro has articulation difficulties in Spanish similar to those observed in English.

Last year Pedro attended kindergarten at that school. Because of Pedro's bilingual background and the results of an oral language proficiency measure that suggest he has limited English proficiency, he first was enrolled in a bilingual kindergarten that stressed English as a second language (ESL) activities as well as a readiness program in Spanish. His academic progress was very slow in this bilingual kindergarten program. This information, together with that from the oral language proficiency measure, suggested the need for an all-English first-grade program. One teacher explained the rationale for this move: "He is limited in both English and Spanish, and academic progress for him is slow and difficult. We felt that a bilingual curriculum was too much for him."

At the time of the child-study team meeting the teachers felt that this intervention was not successful and they sought help in planning Pedro's educational program. Consider this situation from the point of view of the diagnostic information model put forth by Cromwell, Blashfield, and Strauss (1975).

1. What information was considered in placing Pedro into the first grade?

2. What information should have been considered in placing Pedro in the first grade?

3. What justification is there for the intervention strategies used in Pedro's first grade placement?

4. What additional information would you consider necessary in order to make satisfactory recommendations for Pedro's educational program at this point?

5. Comment on the teacher's observation that, because Pedro's progress was too slow, a bilingual curriculum may be too hard for him.

---

[1] Appreciation is expressed to Dr. Juana Canabal-Antokoletz who assumed major responsiblity for developing Simulation Two through Six.

## SIMULATION 3

Andrea (CA-6) is a black first grader at Elmwood Elementary School. She attended kindergarten there last year. Andrea is described by her teachers as friendly and outgoing. She likes school, and her attendance is excellent. Andrea's teachers are concerned about Andrea's highly disruptive classroom behaviors and her failure to acquire basic readiness skills. Her language also is delayed. She has difficulties understanding verbal instructions and has some articulation problems. Andrea is in an open classroom arrangement. She moves to different locations to receive instructions from different teachers in math, reading, and language arts.

A meeting was held in March to discuss Andrea's progress. School personnel and support staff attending the meeting included Andrea's reading, math, and home-room teachers, the school counselor, speech therapist, school nurse, school social worker, and school psychologist. Andrea's mother was invited but she did not come to the meeting.

At the beginning of the meeting, Mrs. W. explained that she is Andrea's maternal grandmother and legal guardian; she has raised her since infancy. Andrea does not know her father. Mrs. W. described Andrea's mother (her daughter) as a limited person who cannot hold a job and who has had inconsistent contacts with Andrea. Mrs. W. also reported that, as a child, Andrea's mother was placed in special education classes and labeled mentally retarded.

As the meeting proceeds you realize that several agendas are present. The official agenda of the meeting is to describe Andrea accurately to determine her special needs, and to offer recommendations for classroom interventions which might be successful in helping Andrea.

Andrea's teachers appear eager to refer Andrea for psychological testing for possible placement in the special education program. Mrs. W. is determined to show the school personnel that she is a competent parent and that Andrea is a normal child. Mrs. W. is determined to prevent the school from placing Andrea in special education classes or giving Andrea any kind of psychological tests. She attributes Andrea's problems to the young teachers' permissive attitudes toward her granddaughter and is willing to advise the teachers regarding discipline strategies to use with Andrea.

1. What kind of additional information (following the ABCD diagnostic intervention strategy model) would be desirable?

2. What possible sources of bias might be present? Identify possible biases associated with (a) the child, (b) teacher, (c) parents, and (d) the school program.

3. Assume that Mrs. W.'s attitude represents an obstacle to a comprehensive assessment of Andrea's functioning. What possible strategies could be used to gain Mrs. W.'s cooperation?

4. Assume that administering standardized tests to Andrea is not possible at this time. How can you help the teachers to gain additional information and plan specific interventions?

## SIMULATION 4

### Role-Playing Activity

Using the case study on Andrea (Simulation 3), create a role-playing situation. Different participants should assume the roles of Mrs. W., the reading teacher (who is very concerned about Andrea's lack of progress), and the school psychologist. Mrs. W. tries to convince the teacher that Andrea needs stricter discipline while the teacher tries to convince Mrs. W. that Andrea's problems have not responded well to stricter discipline and that the teacher needs more information in order to understand and help Andrea. The school psychologist will try to help Mrs. W. and the teacher to arrive at a better understanding of the goals each has and will offer new strategies and encourage others to do so too in order to cope more effectively with the problems.

## SIMULATION 5

**Materials:** Psychological Report

**Assignment:**

Evaluate the attached psychological report. Assess possible sources of bias using the ABCD diagnostic-intervention model.

### Psychological Report

**Name:** Juanita M.
**Age at Testing:** 10-10 **Grade:** 5th

**Tests Administered:**

Wechsler Intelligence Scale for Children-Revised (WISC-R)
Wide Range Achievement Test (WRAT)
Bender Gestalt Visual Motor Test
Sentence Completions
Bilingual Syntax Measure Level II
House-Tree-Person
Kinetic Family Drawing

**Quantitative Results:**

WISC-R

| | English Administration Scaled Score | Spanish Administration Scaled Score |
|---|---|---|
| Information | 5 | 8 |
| Similarities | 7 | 9 |
| Arithmetic | 9 | 9 |
| Vocabulary | 4 | 8 |
| Comprehension | 8 | 10 |
| Digit Span | 7 | 7 |

| | Verbal Scale IQ: 79 | Verbal Scale IQ: 92 |
|---|---|---|
| **Performance Scale** | | |
| Picture Completion | 9 | Not administered in Spanish |
| Picture Arrangement | 10 | |
| Block Design | 10 | |
| Object Assembly | 8 | |
| Coding | 9 | |

Performance Scale IQ: 93
Full scale IQ: 85 (English administration)

| WRAT | Grade Equivalent | Standard Score |
|---|---|---|
| Reading | 4.0 | 89 |
| Spelling | 4.1 | 89 |
| Arithmetic | 4.7 | 96 |

Informal assessment of Juanita's reading and writing skills in Spanish: approximately second grade level. Juanita is able to write sentences expressing complete thoughts. She did not show reversals in her words and her handwriting was intelligible.

## REFERRAL

Juanita was referred by the local referral committee because of concerns about her level of academic achievement and her withdrawn behavior.

## BEHAVIORAL OBSERVATIONS

Juanita is a slender, tall, shy, dark-haired, 10-year-old girl. Her gait is a bit stiff, with toes slightly turned in. She kept her eyes fixed on the examiner throughout testing and had a sober, rather sad-eyed facial expression. She tended to be rather inhibited; therefore, repetition and encouragement were required.

## ASSESSMENT RESULTS

**Language and Communication.** The first examiner used English throughout the testing session. Then Juanita was tested by a bilingual examiner in Spanish, her dominant language, the primary language of the home. The Spanish-speaking examiner reported that Juanita answered most questions with one word sentences although she seemed capable of expressing herself in complete sentences. She also reported her language in Spanish (mastery of syntax, receptive vocabulary) as adequate for her age. Both examiners felt that Juanita shows deficits in English including lack of fluency and vocabulary. In summary, she appears to have an incomplete acquisition of a second language. Juanita may have difficulty understanding classroom instructions at age level. The speech and language teacher assessed Juanita's language and reported that Juanita has some comprehension problems, and her expressive skills are limited. Juanita also has a minor articulation problem. During the testing in English, the examiner noted that Juanita spoke in a soft, almost inaudible voice, and appeared to have a slight lisp.

**Physical/Psychomotor:** Juanita has passed the school's hearing and vision screening. Her mother reported no unusual health or developmental problems. Regarding psychomotor functioning, Juanita made two scorable developmental errors in copying the geometric figures on the *Bender Visual Motor Gestalt Test*. This ranks her performance as falling in the normal range. Her figure drawings also suggested adequate development. Her pencil control (right hand) appeared adequate.

**Emotional/Behavioral/Adaptive:** The Spanish examiner who administered the *Kinetic Family Drawings* found that Juanita visualizes the family as an extended one where grandfather, uncles and aunts are important. The examiner suggested that Juanita's difficulty in learning English may be secondary to her difficulty in becoming accepted socially and feeling that she is part of the English-speaking culture. Though Juanita does not appear to be a severely disturbed child, some depressive feelings were suggested by her drawings and associations.

**Sociological:** Juanita is a native of Venezuela. Her native culture is different from that of the majority in this country and from the dominant culture in the school which she attends. Her social development has been somewhat restricted compared to other girls her age and those included in the test's standardization sample. Positive factors include an intact and extended family, and emphasis on educational achievement in her family.

**Intellectual:** On the *Wechsler Intelligence Scale for Children-Revised*, Juanita's global intellectual functioning was in the average range. When she was unable to respond appropriately in English, questions on the Verbal Scale were repeated in Spanish. This technique gave her the opportunity to demonstrate improved performance on the Vocabulary subtest, and on three other subtests. She showed relative strength in practical judgment and mental computation. On the Performance Scale all five subtest scores fell in the average range. This reinforces the impression that Juanita has average skill in the visual conceptual and perceptual motor areas and nonverbal reasoning. She seems to have some problem with immediate rote memory of nonmeaningful material, such as numbers.

**Educational Performance:** Juanita scored at low fourth-grade level in word recognition and written spelling, and scored at high fourth-grade level in written arithmetic. The Spanish examiner had Juanita write paragraphs and found that her achievement was superior when Spanish is used. Her success on written problems reinforces the impression that she has relative strength in nonverbal learning areas.

## SUMMARY AND IMPLICATIONS

Juanita is an attractive 10-year-old girl. Her global intellectual functioning is in the average range, but her acquisition of English as a second language still is incomplete. The Spanish-speaking examiner noted that Juanita is ready for instruction in English but needs much oral language development in English. There is not a significant discrepancy between Juanita's standard scores on the WRAT and her measured ability on the WISC-R (Spanish and English combined). She is having difficulty adjusting emotionally to a different culture, may be somewhat depressed, but is not severely disturbed.

The two examiners will make arrangements to meet with the mother to interpret the test results. At that time ways will be explored to help Juanita to adapt to her current school and broader cultural life situations. Various alternatives will be explored with the mother, including tutoring and possibly counseling. Juanita needs to be encouraged to participate in activities such as sports and organized arts and crafts. In such situations she could interact with other children without being too pressured to communicate verbally.

After the parent conference, the Local Referral Committee will meet again to review test findings and to consider ways that the school can contribute to Juanita's development.

Answer the following questions:

1. All the tests used in the report are based on standardization populations which are significantly different from Juanita's background. Juanita was born in Venezuela and has been in the U.S. for 18 months. What possible sources of bias may have been introduced by using these tests?
2. What measures were used to minimize bias? Were they sufficient? In spite of the reservations about the standardization populations of the tests used, did the results provide meaningful information about Juanita's functioning? Explain your answer.
3. Even though Juanita came from a different country, the report states that her family values educational achievement. What types of skills may have contributed to Juanita's relative success on the WISC-R?

## SIMULATION 6

You, the school psychologist, are assigned to a high school with a 50% minority enrollment. The administration is concerned about the high dropout rate among minority students. You have been asked to function as a consultant to help the school to encourage minority students to stay in school longer.

Identify possible sources of bias which may be contributing to the high dropout rate of minority students. Formulate potential sources of bias associated with curriculum design and discipline policies. What strategies would you use in order to bring about changes in these areas?

## KEY TO PRETEST

### Short Answer Items

1. Skin color, unusual last name, poor dress, socio-economic status, bilingualism.
2. Teacher interviews, family interviews, school/medical records, child interview, observations, adaptive behavior data.
3. Reliability = Stability of a person's characteristics. Validity = Degree to which a measure provides an accurate description of the behavior it is designed to assess.
4. Norms: Standards against which a child is judged. Nonstandard norms may include parent expectations, other children, teacher expectations.
5. Multiple information sources (e.g., home, school, child), reliable and valid information, effective recommendations from sources familiar with child, responsibilities for interventions willingly assumed by persons with appropriate authority, interventions not limited specifically to child.
6. Insufficient support for special services; discomfort in working with special child populations; dictatorial leadership styles; school isolation.
7. Historical information, current characteristics (child and environment), interventions, follow-up evaluation.
8. Allowing school absences, allowing out-of-school activities to compete with school.
9. Test wiseness: Certain abilities and attitudes prerequisite to successfully taking tests. Components of test wiseness include understanding directions, considering all possible responses before choosing one, nondistractibility (to the extent that response effectiveness is maintained), task involvement, and attentiveness.
10. Motivation, anxiety, cultural background, SES.
11. Skin color, neatness, quality of dress, handicaps, demeanor.
12. Pressure from administrators or colleagues, job security, allegiance to employer.

### True-False Answers

1. True
2. False
6. False
7. True
11. True
12. False

3. False
4. False
5. False
8. False
9. False
10. False
13. False
14. True

## KEY TO POSTTEST

1. (A) *Historical information* helps us to understand important antecedent events in a child's life. (B) *Current characteristics* describe what a child and his environment presently are like. (C) *Interventions* include the processes over which we have control which are intended to effect a positive change in the child or the child's environment. (D) *Prognosis:* This evaluation component estimates the success of interventions.
2. The parents, the child, the examiner, the adequacy of our diagnostic-intervention techniques, and school system policies and procedures.
3. A child may have a general language deficiency or a language difference, because of either exposure to nonstandard dialects or to the exclusive knowledge of a language other than English.
4. Instruction and/or drill and practice in concepts related to basic test-taking skills, such as understanding test directions, multiple choice strategies, concentration skills, and common test concepts (e.g., *same* or *different*), reflective response styles, and eliminating incorrect options.
5. A child is often unavailable for follow-up; administrative policy sometimes precludes time spent on evaluation; the psychologist is often pressured to emphasize treatment, particularly when caseloads are heavy.
6. (a) *Medical Model:* Employs either *ABCD* or *BCD* constructs. Abnormal conditions are detected and treated biologically, based on normative biological standards and dichotomous data (i.e., the presence or absence of symptoms). Has low potential for racial and cultural discrimination. (b) *Social Systems Model:* Employs *BCD* or *BD* constructs and emphasizes learning and displaying socially approved behaviors. It has low potential for racial and cultural discrimination when roles and environments are specified. (c) *Psychoeducational Process Model:* Stresses *BC(D)* relations and emphasizes the assessment and training of processes or abilities usually in a compensatory or remedial program. It has a low source of bias when the child matches the characteristics of the standardization sample. (d) *Task Analysis (Behavioral) Model:* BC(D) type constructs are used to carry out a test-teach-test activity. A child's behavior results from the interaction of a set of enabling behaviors and the requirements of a task. It has a low potential for racial or cultural bias if scores are interpreted in terms of the child's development along a well-sequenced curriculum that provides for the continuous development of enabling behaviors and skills. (e) *Pluralistic Model:* Emphasizes *BD* constructs and uses various techniques in an attempt to discover the talent and potential that are masked by culturally biased and inappropriate measures. Three methods are used or suggested in eliminating bias: the development of culture specific tests, estimates of the child's learning-to-learn ability, and multiple tests norms for various target groups.
7. Improve general perceptual, intellectual, or linguistic processes. Provide more individual attention.
8. The examiner is competent. The sample of behavior is adequate. The characteristics of the child being tested are similar to those children included in the test's standardization sample. These assumptions must be continuously examined for each child with whom we work.

## REFERENCES

Cole, M., & Bruner, J. Cultural differences and inferences about psychological processes. *American Psychologist*, 1971, *26*, 867-76.

Cromwell, R. L., Blashfield, R. K., & Strauss, J. S. Criteria for classification systems. In N. Hobbs (Ed.), *Issues in the classification of children* (Vol. I). San Francisco: Jossey-Bass, 1975.

American Psychological Association, *Ethical Standards for Psychologists*. Washington, D.C.: Author, 1972.

Havighurst, R. Minority subcultures and the law of effect. *American Psychologist*, 1970, 25(4), 313-22.

Newland, T.E. Assumptions underlying psychological testing. In T. Oakland & B.N. Phillips (Eds.), *Assessing minority group children*. New York: Human Sciences Press, 1973.

Mercer, J., & Lewis, J. *System of multicultural pluralistic assessment*. New York: Psychological Corporation, 1978.

Mercer, J., & Ysseldyke, J. Designing diagnostic-intervention programs. In T. Oakland (Ed.), *Psychological and educational assessment of minority children*. New York: Bruner/Mazel, 1977.

Oakland, T. Effects of test-wiseness materials on standardized test performance of preschool disadvantaged children. *Journal of School Psychology*, 1972, *10*, 355-60.

Oakland, T., & Laosa, L. Professional, legislative and judicial influences on psychoeducational assessment practices in schools. In T. Oakland (Ed.), *Psychological and educational assessment of minority children*. New York: Brunner/Mazel, 1977.

## RECOMMENDED READINGS

Three excellent volumes present the findings from the Project on Classification of Exceptional Children which was under Nicholas Hobbs's direction. The thoughts of 93 contributors are presented in the two volumes on *Issues in the Classification of Children*. The diagnostic-intervention model discussed in a preceding subsection of this paper comes from this source. Other articles are of value, too. Hobbs's thoughts, no doubt nurtured by the contributions of others, are presented in *The Futures of Children*. The Project and books are intended to increase the public's understanding of the problems that are associated with the classification and labeling of children; to provide a rationale for public policy and legislation; and to improve practice within various professions. The three books were published in 1975 by Jossey-Bass.

As we know, American education is experiencing a strong trend toward the development of basic skills, recognizing that subsequent behaviors are well developed only if they rest on a firm foundation. This notion applies to one's professional life, too. Many pitfalls could be avoided if a brief moratorium were declared on discussing issues in the assessment of minority children while we read (or reread) *Standards for Educational and Psychological Tests* (American Psychological Association, 1974). By meeting its standards for tests, manuals, and reports, for reliability and validity, and for interpreting scores, we would measurably advance nonbiased assessment. Its most recent revision also recognizes special problems in assessing minority children.

I am fortunate to have participated in various projects directed toward minority children. One of the most satisfying occurred in 1976 when I was asked to prepare a volume discussing the state of the art and science of assessing minority children. Through the contributions of Drs. Bernal, Laosa, Matuszek, Mercer, Tucker, Ysseldyke, and others, the book, *Psychological and Educational Assessment of Minority Children* (Bruner/Mazel, 1977) emerged under my editorship. It contains discussions of historical antecedents and current issues; professional, legislative, and judicial influences on psychoeducational assessment; using tests in nondiscriminatory assessment; designing diagnostic-intervention programs; and operationalizing a diagnostic intervention process. The appendices present many of the original federal guidelines impacting on nonbiased assessment (e.g., Office for Civil Rights memorandum and the remedies for the Lau v. Nichols decision). The annotated bibliography of 27 language-dominance measures also has been helpful to practitioners.

# NONBIASED ASSESSMENT: LEGAL PRINCIPLES

**Donald Bersoff**
**University of Maryland**

# Table of Contents

Introduction and Objectives ........ 19

Pretest ........ 20

The Legal System's Conception of Nondiscriminatory Assessment ........ 21
  1. Case Law ........ 21
    a. *Hobson v. Hobson* ........ 21
    b. The Case of *Larry P.* ........ 22
    c. A Critique of the Case Law ........ 24
    d. Interim Summary ........ 25
  2. Legislation ........ 25
    a. Public Law 94-142 ........ 25
    b. Rehabilitation Act of 1973 ........ 26
    c. Interim Summary ........ 26

Procedural Requirements in the Performance of Nondiscriminatory Assessment ........ 26
  1. Notice and Consent ........ 27
    a. The Meaning of Consent ........ 27
    b. When Consent Is Required ........ 28
  2. Content of Notice and Consent Forms ........ 28
  3. Interim Summary ........ 29

Conclusion ........ 29

Theory Into Practice: Simulations ........ 30

Pretest Key ........ 31

References ........ 31

Table of Cases ........ 31

Recommended Reading ........ 32

## Introduction

Every person in the United States probably has been affected by tests in some way. An estimated 250 million standardized tests of academic ability, perceptual and motor skills, emotional and social characteristics, and vocational interests and talent are used each year in educational settings (Holmen & Docter, 1972). Test results have been used as both facilitative devices—to admit, advance, and employ—and exclusionary devices—to segregate, institutionalize, track, and deny access to desired goals. Although testing has been criticized by social, political, and psychological commentators for six decades, the legal system seriously began to examine the issue of nondiscriminatory assessment only in the last 10 years. In part, this recent interest may be explained by the Supreme Court's mandate in *Brown v. Board of Education* (1954) to desegregate public schools. Civil rights advocates began to view educational and psychological tests as tools to hinder integration and, more broadly, as discriminatory instruments denying racial and ethnic minorities the full realization of their constitutional rights. As a result, since the mid-1960s, an explosion of litigation and legislation has affected the administration, interpretation, and use of educational and psychological tests.

In this section some fundamental legal principles which are derived from major cases and statutes on nondiscriminatory assessment are examined and discussed. It would be helpful to know the legal history of school-testing litigation, the legal theories under which tests have been challenged, all current regulations and statutes that control the administration of tests in public school, and issues subsumed under these topics. Because such a comprehensive survey, in the context of this module, would be overwhelming (for such reviews see Bersoff, 1979; Oakland & Laosa, 1977), two basic questions will be discussed: (a) How does the legal system define nondiscriminatory assessment? (b) What kinds of procedural requirements does the legal system impose on school systems to enhance the possibility that minorities will not suffer discrimination as the result of taking tests?

The material is structured as follows to help answer the two questions: (a) A series of questions in the format of a pretest alerts readers to the major concepts that follow; (b) An extended content section discusses information that is relevant to the topic. Interspersed within the section are discrete activities, which readers are expected to complete. (c) Simulation exercises are provided to help to integrate the theory into practice and to serve as an assessment of skills and knowledge gained in this section.

## PRETEST

Matching, Item 1.

1. Five important cases affecting nondiscriminatory assessment are listed in (a). Some major principles from these decisions are presented in (b). Next to each case write in the number of the major principle or decision that comes from that case.

(a) Cases

______ Hobson v. Hansen
______ Larry P. v. Riles
______ Brown v. Board of Education
______ Washington v. Davis
______ Griggs v. Duke Power Co.

(b) Principles:

(i) Placement of black children in EMR classes when the primary determinant of placement is an individually administered intelligence test score violates the equal protection clause of the U.S. Constitution, P.L. 94-142, and Section 504 of the Rehabilitation Act.

(ii) Segregation of black children in separate schools violates the equal protection clause.

(iii) Use of intelligence tests that are fair in form but have a discriminatory effect (i.e., lead to disproportional selection to the disadvantage of black persons) violates the 1964 Civil Rights Act.

(iv) Tracking students in ability groups on the basis of racially and culturally discriminatory standardized group tests violates the equal protection clause.

(v) When plaintiffs prove that the use of intelligence tests only has a discriminatory effect without also proving that their use was intentionally designed to discriminate against minorities the equal protection clause is not violated.

True-False, Items 2-10

Write your answer next to each statement. If it is false, state what you think is the correct principle.

2. Classification practices that result in racial disparity violate the equal protection clause.
3. The rights of black children under the 1964 Civil Rights Act and the U.S. Constitution are essentially equivalent.
4. If tests are the best means available to assess children, their use does not violate the equal protection clause even if disproportionately more black children are placed in EMR classes.
5. Under P.L. 94-142 parents may contest the school psychologist's evaluation and recommendations in an impartial hearing.
6. School psychologists may not administer tests unless parents consent.
7. Only assessment instruments that measure innate capacity should be considered nondiscriminatory.
8. Parents who are mentally retarded are legally incompetent to consent to a psychological evaluation.
9. If consent forms are written in plain, easy to read English, they will meet the requirement in P.L. 94-142 that such forms must be understandable to parents.
10. School systems must notify parents before they may reevaluate children already placed in special education programs.

Short Answers, Items 11-13

Under each of the following 3 questions provide the information requested in a written statement:

11. State under what circumstances Sec. 504 of the 1973 Rehabilitation Act, but not P.L. 94-142, will control the administration of psychological tests to minority children.

12. List the three basic pieces of information that must be given to parents prior to the administration of a psychological evaluation.

13. P.L. 94-142 requires a "comprehensive evaluation" prior to placement in a special education program. Define the meaning of that term.

Multiple Choice, Items 14-18. Circle the correct answer

14. The equal protection clause of the Fourteenth Amendment prevents discrimination by:
    a. States
    b. Private individuals
    c. Federal government
    d. Any institution receiving federal financial assistance
    e. All of the above

15. To meet the requirements of the Constitution and federal law, a test must have, at a minimum, which of the following validity coefficients to be considered nondiscriminatory:
    a. .20
    b. .50
    c. .75
    d. .90
    e. None of the above

16. Which one of the following is *not* a requirement of P.L. 94-142's provisions on the assessment of handicapped children:
    a. Tests must be administered in the child's native language
    b. Intelligence tests must not be the sole determinant of placement
    c. All evaluations must be performed by a credentialed school psychologist
    d. Tests must be validated for the purpose for which they are used
    e. None of the above

17. Under federal law parents must give their consent to:
    a. All individual evaluations by a school psychologist
    b. All individual evaluations by any member of the psychoeducational evaluation team
    c. Only to initial evaluations and to triennial reevaluations
    d. Only to evaluations performed prior to initial placement in a special education program
    e. All assessment used for identification (screening) and potential placement

18. It is a violation of federal law and decisions in several cases to administer which of the following tests to minority children:
    a. Standardized group intelligence tests
    b. Standardized individual intelligence tests
    c. Projective tests
    d. Personality inventories
    e. None of the above

This last series of 5 questions relate specifically to your practice as a school psychologist. On a scale of 1-5 rate your performance on each of the following descriptive statements. Rate yourself "1" if your answer is "Never" and "5" if your answer is "Always." Place your rating on the line next to the statement. A rating of 1 is not necessarily negative.

19. ______ The primary determinant I use for diagnosing a child as educable mentally retarded is an individually administered intelligence scale

20. ______ In evaluating a child from an ethnic minority I assess whether the child's primary language is English

21. ______ I study the test manual to see if the publisher has reported validity coefficients for the purpose for which I will use the test

22. ______ Before I evaluate any child prior to placement I make sure that there is a form, signed by the parent or other appropriate person, consenting to the evaluation

23. ______ Before I evaluate any child prior to placement I have described each of the assessment procedures I plan to use to the parent

## The Legal System's Conception of Nondiscriminatory Assessment

### 1. Case Law

For a dozen years after *Brown v. Board of Education* (1954) many school systems attempted to forestall the process of desegregation by introducing innovative mechanisms that would preclude black children from attending previously all-white schools. Many of these tactics relied heavily on the use of intelligence and achievement tests. For example, in one major southern city, no black child was permitted to transfer to a "white" school unless his/her grade-level score on an ability test was at least equal to the class average in the school to which the transfer was requested. This and other dilatory mechanisms were challenged in the federal courts by minority plaintiffs and eventually struck down as unconstitutional. However, no case attacked the validity of the tests. The only concern of the judiciary at the time was whether standardized tests were administered only to blacks or were used to make decisions solely on racial grounds. During the first decade following *Brown*, test administration was free from judicial scrutiny. In the early 1960s, when the courts were attempting simply to begin the process of desegregation, charges that tests themselves were racially and culturally biased were not yet heard.

**a. Hobson v. Hansen**

Against this background a federal district court in Washington, D.C., heard and ultimately decided *Hobson v. Hansen* (1967), the first case in which a court directly ventured into the complexities of the testing controversy. The case generally concerned the legality of intra-school district disparities in financial and educational resources as a result of which white children were receiving better education and more monetary support. At the heart of the matter, however, was the disproportionate number of black children in lower tracks, most of whom had been placed in the tracks on the basis of standardized group tests.

Despite the fact that the District of Columbia had instituted ability grouping in a genuine attempt to remedy the severe academic deficiencies of black children, the Court ultimately condemned the tracking system because it found significant racial disproportionality among the children assigned to the different tracks. "As a general rule," the Court said, "in those schools with a significant number of white and Negro students a higher proportion of Negroes will go into the Special Academic (EMR) Track than will white students" (p. 456). The Court was also concerned with the showing that placement in tracks was determined primarily on the basis of standardized test results. Although the stated criteria for entrance to a track included teacher and counselor evaluations of maturity, stability, physical condition, and grades, the Court found that "testing looms as the most important consideration in making track assignments" (p. 475). Thus, it was the dual-pronged finding of disproportional placement determined primarily by reliance on test scores that triggered the Court's extensive inquiry into the nature and limitations of standardized group tests.

The legal question facing the Court was whether racial disparities among the tracks violated the equal protection clause of the Fourteenth Amendment which prevents states (school systems are considered to be arms of the state for constitutional purposes) from denying "to any person within its jurisdiction the equal protection of the laws."[1] The equal protection clause is not an absolute barrier to classification, and not all classifications resulting in disparity are unconstitutional. If the classification is reasonably related to the purposes of governmental activity and is performed fairly, the fact that persons are treated differently is not necessarily a violation of the constitution.

Thus, the Court in *Hobson* undertook to determine if ability grouping, a form of classificatory activity, has a rational basis. The Court asserted that discrimination on the basis of ability could be defended only if judgments on ability were based on measures that assessed children's innate endowment or capacity to learn, not their present levels of skills. If so, ability grouping would be viewed as "reasonably related to the purposes of education" (p. 512). The law has a special concern for victims with a long history of purposeful and malicious discrimination, as in the case of racial minorities. The defense of practices that appear to continue to increase discrimination is difficult. Thus, the school system was asked to explain why black (and poor) children disproportionately populated the lower tracks.

The only explanation the Court would allow for the racial disparities was that the tests upon which the classifications were based accurately reflected students' innate abilities. The Court concluded that the evaluation instruments did not pass that test:

> While the government may classify persons and thereby effect disparities in treatment, those included within or excluded from the respective classes should be those for whom the inclusion or exclusion is appropriate; otherwise the classification risks becoming wholly irrational and thus unconstitutionally discriminatory. It is in this regard that the track system is fatally defective, because for many students placement is based on traits other than those on which the classification purports to be based. (p. 513)

The Court next wrote the statements that were to have a profound effect on the use of educational and psychological tests to this day:

> The evidence shows that the method by which track assignments are made depends essentially on standardized aptitude tests which, although given on a system-wide basis, are completely inappropriate for use with a large segment of the student body. Because tests are standardized primarily on and are relevant to a white middle class group of students, they produce inaccurate and misleading test scores when given to lower class and Negro students. . . . (T)hese students are in reality being classified . . . on factors which have nothing to do with innate ability. (p. 514)

---

*Task 1*

*From a school psychologist's point of view, write a brief critique of the court's position that tests are valid only if they measure innate ability*

---

*Hobson* represents the condemnation of rigid, poorly conceived classification practices that negatively affect the educational opportunities of minority children. The Court's major concern was not the tests but inflexible ability grouping, the tracking system's stigmatizing effect on blacks and its failure to provide sufficient resources to students in the lower tracks. EMR programs were viewed as relegating students to permanent inequality. An examination of ability tests was restricted to their use as sole or primary decision-making devices to justify placement.

One significant finding by the Court was that reliance on group measures contributed to the misclassification of approximately 820 of 1272 students. Evidence of misclassification was provided by the school system itself. In 1965, two years prior to the *Hobson* decision, the superintendent of schools had ordered that no student be assigned to the EMR track without an evaluation (usually an individual test) by a psychologist. When clinicians then reassessed the EMR children, they concluded that almost two-thirds were not retarded.

In that light, *Hobson* could be read as support for the use of individual tests by school and clinical psychologists.[2] However, soon after the case was decided, minority plaintiffs in California and several southwestern states began a round of post-*Hobson* cases of significant dimension. Despite *Hobson's* implicit approval of individual testing, new cases began to attack the stately, revered, and venerated devices against which all other tests were measured: the individually administered intelligence scales like the *Stanford-Binet* and *WISC*.

### b. The Case of Larry P.

By far the most important of the post-*Hobson* cases is *Larry P. v. Riles* (1972, 1974, 1979), perhaps better known as *Larry P.* The case presents severe threats to the continued administration of individual intelligence tests, particularly with minority children, and raises significant statutory and constitutional issues that will affect future litigation on psychological assessment.

**Phase One.** *Larry P.* began in 1971 as a class action suit initiated on behalf of black children who were placed in EMR classes ostensibly because their IQ scores were lower than 75 on state-approved intelligence tests, predominantly the *Stanford-Binet* and *WISC*. The plaintiffs, claiming that they were not mentally retarded and that the tests used to place them were culturally biased, alleged that the resultant misclassification violated their rights under the Constitution's equal protection clause. Thus, the Court was asked to grant an injunction temporarily prohibiting the defendant, the San Francisco School System, from administering IQ tests to determine the placement of black children in EMR classes until a full trial could be held to decide the merits of the complaint.

In supporting their claim of misclassification, the plaintiffs presented affidavits from several black psychologists showing that the children scored above the cut-off point for EMR placement after retesting in which special attempts were made to establish rapport, items were reworded, and scoring procedures were changed. Although plaintiffs normally are required to show that the defendant's actions were arbitrary and irrational, the plaintiffs requested instead that the school system *prove* that its classification process was reasonable.

The Court required the plaintiffs to make two showings before it would order the school system to prove that its practices were rational: (a) that racial imbalance existed in the composition of EMR classes; (b) that the primary determinant for placement in those classes was test scores derived from the challenged intelligence tests. Although blacks constituted only 28 per cent of the students in the San Francisco School System, 66 per cent of the EMR students were black. Similarly, blacks comprised 9 per cent of the

California school population and 27 per cent of EMR school children. Granted that placement in EMR classes was based on intelligence test results and not race, the plaintiffs alleged that this method of classification led to a disproportionate impact on black children. Thus, the data regarding racial imbalance for EMR children were undisputed and unequivocal. Satisfied that the first prong of the test was met, the court proceeded to analyze if the primary basis for EMR placement was the IQ test.

In California's Education Code, school psychologists were required to conduct comprehensive evaluations of children's developmental history, cultural background, school achievement, adaptive behavior, and intelligence. The Court found that the district placed primary emphasis on IQ scores in making assignments to EMR classes and it pointed to a California statute that required other data to substantiate the IQ.

The Court, knowing that racial imbalance existed and that EMR placement was determined primarily by IQ tests, asked the defendants to justify their use of IQ scores and to "demonstrate the rational connection between the tests and the purpose for which they are allegedly used" (p. 1311). The school system responded by candidly agreeing that the tests were racially and culturally biased and justified their continued use as the best means available to classify students as retarded; suitable options did not exist.

The San Francisco School District also sought to defend the disparities by claiming that racial imbalance was the result of (a) white parents placing their retarded children in private schools more frequently than did black parents, (b) the district's prior practices of locating EMR classes in predominantly black schools, and (c) black parents tending to be poorer and, thus, producing children with a higher incidence of retardation owing to nutritional deficiencies. These defenses were rejected because, in the Court's judgment, the school system produced no hard evidence to support them.

Thus, the Court in 1972 held the school district to be in violation of the equal protection clause because it failed to demonstrate that "IQ tests are rationally related to the purpose of segregating students according to their ability to learn in regular classes, at least insofar as those tests are applied to black students" (p. 1313). The Court issued the preliminary injunction prohibiting the placement of black children in EMR classes on the basis of intelligence test results if the consequence was racial imbalance in such classes.

Three events followed the Court's decision: (a) An appellate tribunal in 1974 affirmed the lower court's order. (b) The court then approved the plaintiff's motion to broaden the injunction to prohibit the administration of individual intelligence tests to all black children in the state. (c) Finally, in 1975, California itself decided to broaden the ban to prohibit the use of IQ tests for placing any California children in EMR classes regardless of their race.

Phase One of *Larry P.* was ended. Phase Two, the trial on the substantive issues, began in October 1977 and ended in mid-1978. The Court published its opinion in October 1979, making permanent the preliminary injunction it had granted in 1972. Phase Three, the appeal by the State of California, began in January 1980 and no definite termination date is in sight.

**Intervening Events.** Between Phases One and Two the Supreme Court and Congress took significant action which had important effects on the outcome of Phase Two. In 1976 the Supreme Court decided *Washington v. Davis*, a case that arose when two black applicants for positions as police officers filed suit contending that the written personnel test used by the police department excluded disproportionately high numbers of black applicants. The federal court of appeals applied the reasoning of *Griggs v. Duke Power Co.* (1971) to resolve the plaintiffs' constitutional argument that the use of the test invidiously discriminated against blacks. The Supreme Court, however, reversed the lower court's decision, rejecting the contention that the *constitutional* standard for adjudicating claims of racial discrimination was identical with the *statutory* standard under the 1964 Civil Rights Act: "(O)ur cases have not embraced the proposition that a law or other official act, without regard to whether it reflects a racially discriminatory purpose, is unconstitutional solely because it has a racially disproportionate impact" (*Washington v. Davis*, 1976, p. 239). Thus, the Court declined to apply the more rigorous standard of the Civil Rights Act to the constitutional questions.

The plaintiffs' task during Phase Two was made more difficult, consequently, because they were required to provide evidence of intent to discriminate and not merely of a discriminatory effect.

Congressional enactment of Sec. 504 of the *Rehabilitation Act* (1973), which prohibits discrimination against handicapped persons in any institution receiving federal financial assistance, and Public Law 94-142, The Education for All Handicapped Children Act of 1975, both have a bearing on *Larry P.* The two statutes and their regulations prohibit the use of assessment instruments that are racially or culturally discriminatory; they require the tests used with children suspected of being handicapped to be validated for the specific purposes for which they are used.

**Phase Two.** The plaintiffs filed and amended their complaint in 1977 alleging that the use of individual intelligence tests for EMR placement violated Title VI of the 1964 Civil Rights Act (in which only discriminatory effect, not intent, would have to be proven) as well as the equal protection clause. Later that year the plaintiffs filed a second amended complaint alleging violation of Public Law 94-142. The U.S. Department of Justice, participating as a friend of the court, sided with the plaintiffs, and also asserted that the state's conduct violated Public Law 94-142 and Sec. 504 of the Rehabilitation Act.

Judge Peckham decided in favor of the plaintiffs on both statutory and constitutional grounds. The defendants were permanently prohibited "from utilizing, permitting the use of, or approving the use of any standardized intelligence tests . . . for the identification of black EMR children or their placement into EMR classes, without securing prior approval by this court" (P. v. Riles, 1979, p. 104).

The Court began its analysis by evaluating the defendants' conduct under Title VI, Rehabilitation Act, and P.L. 94-142. The Court, finding no data validating the use of IQ tests for EMR placement, held that EMR placement mechanisms that lead to a disproportionate number of black children being deprived of a meaningful education are in violation of Title VI of the Civil Rights Act.

The Court's primary focus was on the nondiscriminatory assessment provisions of Sec. 504 and P.L. 94-142, particularly that part of the regulations requiring assessment instruments to be "validated for the specific purpose for which they are used." The Court's interpretation of these provisions is crucial to its decision and breaks new ground because it is the first case that attempts to establish validation criteria for tests used for EMR placement. Relying on Griggs v. Duke Power Co. (1971), the Court said that the state first must show that the test has a direct relationship to the position for which the test is required. If the relation can be shown, the plaintiffs then must submit evidence to show that alternative selection procedures exist that would serve the employers' purposes as well without producing discriminatory effects.

Judge Peckham held the state accountable for proving that the tests used for placement had been validated for black children and for showing that the tests were valid for selecting children who are unable to profit from instruction in regular classrooms with remedial instruction and, thus, belonged in—to use the Court's term—isolated, dead-end, stigmatizing EMR programs. According to the Court, the State did not provide suitable kinds of validation.

> (D)efendants must come forward and show that . . . (the tests) have been validated for each minority group with which they are used . . . . This minimal burden has not been met for diagnosing the kind of mental retardation justifying the E.M.R. placement. (p. 69)

The few studies that had been brought to the Court's attention were not considered relevant. The Court rejected validity studies in which IQ scores correlated with college grades or other achievement tests. It was satisfied only with research relating IQ scores of black children with classroom grades, although the latter were admittedly subjective. The one relevant study which was cited yielded correlations between IQ scores and grades of .25 for white children and only .14 for blacks. The expert who testified about the study concluded that the WISC had "little or no validity for predicting the scholastic performance of black or brown children" (p. 71). Thus, the Court concluded that

> The IQ tests are differentially valid for black and white children. . . . Differential validity means that more errors will be made for black children than whites, and that is unacceptable. (p. 71-72)

The Court found that alternative mechanisms existed for determining placement in EMR classes. During the statewide IQ moratorium, adequate assessments were made without IQ tests, and fewer misplacements occurred. In fact, the Court found that more time and care had been taken during this period in placing children in EMR classes than when IQ tests were administered. Nevertheless, the Court warned, alternatives themselves to IQ tests had not been validated, and disproportionate placement, although less egregious than in the pre-1975 era, was still present. Continued use of tests is needed for "the development of curricula that respond to specific educational needs" (p. 74), not for the purpose of labeling children as retarded. Thus, given the functional exclusion of black children under Title VI and the failure to meet its burdens under Section 504 and P.L. 95-142, the Court found the defendants in violation of the two statutes. Judge Peckham also ruled that the plaintiffs warranted relief under the Fourteenth Amendment to the Constitution. The primary question of judging intentional discrimination was decided by requiring the plaintiffs to show an intent to segregate children into classes for the educably retarded. This decision to define purposeful discrimination as the intent to segregate minority children into special classes laid the groundwork for vindicating the plaintiffs' claims. In the end, the Court was satisfied that the plaintiffs had met this burden. The Court found EMR programs to be horribly ineffective and blacks to be substantially overrepresnted in them. After concluding again that IQ tests were the primary determinant in EMR placement, the Court found the tests to be racially and culturally biased. It cited problems in the standardization of the tests, in the exclusive use of standard English, and in items not reflecting acceptable cultural differences.

The Court concluded that the "Defendants can establish no compelling state interest in the use of the IQ tests nor in the maintenance of EMR classes with overwhelming disproportions of black enrollment" (p. 94).

After finding for the plaintiffs under federal law, the Constitution, and the California Constitution, the Court proposed the following remedies: not to use standardized intelligence tests to identify black children for EMR placement without first securing approval from the Court; to eliminate overrepresentation in EMR classes by race and ethnicity; to reevaluate all black children labeled EMR without resort to standardized intelligence tests not approved by the Court; and to draft individual education plans designed to return all incorrectly identified children to regular classrooms.

**Phase Three.** The recent appeal of Phase Two and the introduction of Phase Three necessitates caution in interpreting the present findings and specification of remedies. Clear guidelines regarding nondiscriminatory assessment from the judiciary still are lacking.

### *Task 2*

*(You may do this task now or after you finish reading the content material.)*

*Locate the statute in your state that prescribes the evaluation procedures for placement in programs for the handicapped. The statute is located in the codification of your state's laws and can be found in any public or law library. By searching through the index you should find the pertinent provisions (or ask the librarian or a law student to help you). After you find the appropriate statute and read through it, write out the criteria for placement; pay particular attention to the emphasis on the use of IQ scores.*

### c. A Critique of the Case Law

In Task 1 you wrote brief critique of the Court's position in *Hobson v. Hansen* (1967) that tests are valid only if they measure innate ability. Similar reasoning was used to condemn individual tests in Phase I of P. v. Riles (1972).[3] Although the specific outcomes in these cases may be applauded, given their facts, the courts showed little knowledge or understanding of measurement theory. Equating nondiscriminatory assessment with instruments that tap inherited ability alone gravely misconceives the nature of tests. No psychologist who has written on the subject believes that psychometric devices measure hereditary endowment solely.

"There are no measures of innate capacity" (Cleary, Humphreys, Kendrick, & Wesman, 1975, p. 17). "Intelligence as measured is not a capacity. It is a behavioral trait and one highly dependent on past learning, whether the test is a standard test of intelligence or a 'culture-fair' test" (Cleary et al., 1975, p. 22). The consensus is that aptitude scales measure developed abilities that reflect the interaction between an individual's experiences and his or her innate talent. No group, whether minority psychologist (e.g., Bernal, 1975; Jackson, 1975) or committed hereditarian (e.g., Jensen, 1969) considers intelligence test scores to reflect only innate ability. The argument among psychologists and other scientists is over the relative weights to assign to hereditary and environmental factors. The Court in *P. v. Riles* (1979; Phase Two) recognized the error in *Hobson* and concluded that intelligence tests merely measure ability, not capacity.

By definition, achievement and intelligence tests always will fail to meet *Hobson's* demand for assessment devices devoid of environmental bias. Given what they purport to measure, tests inevitably reflect the social setting in which the test-taker lives. "All behavior is affected by the cultural milieu in which the individual is reared and since psychological tests are but samples of behavior, cultural influences will and should be reflected in test performance. It is therefore futile to try and devise a test that is free from cultural influences" (Anastasi, 1976, p. 345).

If meeting the *Hobson* innate capacity test is impossible evidence demonstrating a minimal to substantial relationship between an assessment practice (means) and labeling or classification (end) of children may be almost as difficult to provide. The issue, when translated into psychologists' terms, becomes, "What level of validity is required to support a placement decision based on an assessment instrument?" Unfortunately, the courts have provided no general answer to how high such a validity coefficient should be.

The law may invoke at least three different standards for evaluating validity: Does it predict significantly better than chance? Does it significantly reduce classification errors? Does it correctly identify children who cannot learn? In 1972, a federal appellate tribunal barred the use of the Graduate Record Examination for the purpose of hiring teachers in a southern school district because the validity coefficient was too low to prevent "the elimination of *some* good teachers" (*Armstead v. Starkville Municipal Separate School District*, 1972, p. 280; emphasis added). Apparently, then, even a few false negatives render a test invalid or, in legal terms, arbitrary and irrational. A federal district court in Pennsylvania ruled unconstitutional the use of personality questionnaires by a school district to identify potential drug abusers among junior high school students. It was concerned, in part, that there would be too many false positives owing to the psychometric deficiencies of the instrument. The Court observed, "when a program talks about labeling someone as a particular type and such a label could remain with him for the remainder of his life, the margin of error must be almost nil" (*Merriken v. Cressman*, 1973, p. 920). "Nil" implies nearly perfect correlation. Few, if any, assessment instruments yield reliability, much less validity, coefficients above .95.

Thus, validity coefficients that school psychologists might find acceptable may not pass judicial muster. Measurement experts may consider the courts' understanding of the levels of validity required to meet the law's rational basis principle naive, but the validity of psychological tests used to classify racial and ethnic minorities clearly must meet a relatively high, albeit unspecified, standard.

### d. Interim Summary

We are now about midway in considering our first major question: What is the legal definition of nondiscriminatory assessment? You may have gathered that there is no simple answer. From the case law, however, we can assert some partial definitions, although not all the following principles are necessarily settled law.

1. Assessment practices must be reasonable and rationally related to the goal of classification and placement.
2. Assessment that does not result in a disproportionate and disadvantageous effect on minority groups may be considered nondiscriminatory.
3. If racial disparities in placement occur, assessment may be nondiscriminatory if the effect is not explicitly intentional and the tests that produce the disparities do not have the natural, foreseeable, or anticipated consequence of creating the disproportionality.
4. Assessment devices must be valid for each discrete minority group with which they are used.
5. Assessment results must produce equal validity coefficients for all groups on relevant predictive criteria (e.g., actual classroom performance).

## 2. Legislation

Thus far we have examined only one of the two major sources of the legal definition of nondiscriminatory assessment. Legislatures as well as courts develop rules of conduct. Whereas courts must wait for litigants to present legally cognizable issues, legislatures and government agencies may pass laws and promulgate regulations at any time they believe (a) problems need a broader solution than courts can provide, (b) judicial decisions too narrowly define rights, or (c) they disagree with the courts and wish to undo their decisions. As long as lawmakers stay within the confines of the Constitution they are free to enact statutes and rules for any of these purposes. This process is evident with regard to psychometric evaluations. In the last half-dozen years both state and federal governments have drafted comprehensive legislation affecting educational and psychological assessment. The next subsections review the two most important federal statutes.

### a. The Education for All Handicapped Children Act of 1975 (PL 94-142)

Passed by Congress in 1975, PL 94-142 is essentially a grant-giving statute providing financial support for special education and related services to state and local education agencies if they meet eligibility requirements. Earlier legislation (PL 93-380) specified that any assessment devices used "for the purposes of classification and placement of handicapped children will be selected and administered so as not to be racially and culturally discriminatory." PL 94-142 and its regulations published in 1977 reaffirmed this nondiscriminatory evaluation,[4] mandate and fleshed out its meaning. Sec. 121a.532 of the regulations states,

> (a) Tests and other evaluation materials:
>
> (1) Are provided and administered in the child's native language or other mode of communication. . . .
>
> (2) Have been validated for the specific purpose for which they are used; and
>
> (3) Are administered by trained personnel in conformance with the instructions provided by their producer. . . .
>
> la-
>
> tions, Sec. 121a.532)

The most ambiguous of these provisions is (a2). The regulations require test *validation* but not test *validity*. Even if one infers that both are necessary, there is no indication of the level of validity to which a test must conform. As noted previously, there are few clear-cut judicial or statutory guidelines for the standards of validity in school testing. To date, only *P. v. Riles* (1979) has given us some indication of what this requirement means.

Other provisions also affect psychological and educational assessment. Children with sensory, manual, or speaking impairments are to be given tests that reflect genuine deficits in aptitude or achievement, not their impairments. Further, all assessment is to be comprehensive, multifaceted, and multidisciplinary. Evaluations for placement must be conducted by persons from education, medicine, and psychology who assess children "in all areas related to the suspected disability, including, where appropriate, health, vision, hearing, social and emotional status, general intelligence, academic performance, communicative status, and motor abilities." Tests cannot be used which are "merely . . . designed to provide a single general intelligence quotient" nor can one single procedure be "used as the sole criterion for determining an appropriate educational program for a child." In making placement decisions the school is required to "draw upon information from a variety of sources, including aptitude and achievement tests, teacher recommendations, physical condition, social and cultural background and adaptive behavior." Thus, PL 94-142 clearly states that assessment and placement decisions are the responsibility of a multidisciplinary team and not of a school psychologist acting alone. This diffusion of responsibility presumably reduces individual bias and broadens accountability.

To insure that all these provisions are used, both the statute and the regulations provide mechanisms enabling parents "to present complaints with respect to any matter relating to the identification, evaluation, or educational placement" of their children. The complaints are presented at an impartial administrative hearing in which parents have the right to compel the attendance of, and to cross-examine, witnesses who participated in the assessment and programming decisions. Thus, psychologists are vulnerable to intense scrutiny of their credentials and performances, including the reliability and validity of the evaluation measures they employ, the interpretations they make from the information gathered, or the recommendations they offer as a result of their evaluations.

**b. Rehabilitation Act of 1973**

In this multipurpose law to promote the education, employment, and training of handicapped persons, Congress made the following declaration in the final provision:

> No otherwise qualified handicapped individual in the United States . . . shall, solely by reason of his handicap, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving federal financial assistance. (Sec. 504)

This section thus represents the first federal civil rights law protecting the rights of handicapped persons and reflects a national commitment to end discrimination on the basis of handicap.

The language of Sec. 504 is almost identical with that of the 1964 Civil Rights Act and can be interpreted to be just as encompassing. Unlike PL 94-142, however, the protection of Sec. 504 is not triggered by the receipt of funds under a specific statute but by the presence of handicapped persons in all institutions receiving federal financial assistance. Thus, any school system, public or private, receiving federal monies for any program or activity whatsoever is bound by its mandates.

In mid-1977, after much delay and many drafts, the Office for Civil Rights (D/HEW) set forth regulations pertaining to the evaluation of children who are suspected of being handicapped. The language of those provisions is almost identical with that which now appears in the regulations for PL 94-142 requiring preplacement evaluations, validated tests, multidisciplinary comprehensive assessment, and periodic reevaluations; hence these regulations are not repeated here.

**c. Interim Summary**

The definition of nondiscriminatory assessment developed by Congress and D/HEW is more general and less rigorous than that given in judicial decisions. The major components of the D/HEW and Congressional definition may be summarized as follows:

1. Tests should be administered in a language a child can understand.
2. Tests should be administered in a standardized manner by qualified personnel.
3. Assessment must be comprehensive, measuring all pertinent aspects of a child's functioning, including adaptive behavior as well as educational and intellectual abilities.
4. No placement decisions can be based on a single measure, particularly an intelligence test.
5. All evaluation and placement decisions must be made by a multidisciplinary group of professionals.

Two comments about this set of operational definitions are in order: (a) It gives no indication of what particular tests meet the criteria of nondiscriminatory assessment. (b) It is still the responsibility of the courts to construe these rules. Despite the law's seeming lack of rigor, the courts may require tests to have the same attributes as those specified in the cases we have reviewed.

In general, the major principle is that nondiscriminatory assessment is a procedure that results in a lack of significant racial and ethnic disproportionality. From case law, the various statutes protecting handicapped students, and the interpretive guidelines developed by governmental agencies, there comes the notion that legal inquiry is triggered by the overrepresentation of minorities in programs that are more stigmatizing, restrictive, and inflexible than those in which white children are placed.

### Procedural Requirements in the Performance of Nondiscriminatory Assessment

Justice Frankfurter remarked that the history of constitutional guarantees was largely to be found in the history of procedure; that is, the process by which a government affords its citizens substantive rights may determine whether those rights are, in reality, exercised. For example, a rule stating that criminal defendants must pay a $500 filing fee within three days of conviction if they wish to appeal essentially nullifies a poor defendant's ability to appeal. Similarly, the right to nondiscriminatory assessment may be hollow if parents never are informed of their children's pending evaluations, are barred from reviewing test findings, cannot secure independent evaluations to substantiate or refute the schools' assessments, or have no opportunity to challenge the results of the assessment in a hearing before an impartial adjudicator. These provisions form a part of the general assessment process and are not limited to nondiscriminatory assessment. Because the courts and federal government find it difficult to define precisely the substantive right itself, they have focused on procedural guarantees as one important way to help insure the psychoeducational assessment will not have a discriminatory impact on minorities.

A recent decision is a useful example. In *Lora v. Board of Education of City of New York* (1978), black and Hispanic emotionally disturbed students claimed that their disproportionate referral and assignment to special day schools that are removed from the mainstream of regular education was discriminatory. Because of vague and subjective criteria for identification, evaluation, and placement, plaintiffs charged that they were treated less favorably than white students who were more likely to be placed in classes for the emotionally handicapped and not in separate schools. Part of the plaintiffs' complaint was that inadequate and improper evaluations led to these disparities.

The court spent a great deal of its 90-page opinion recounting the process by which children were evaluated. Candidates for placement in special day schools were assessed by "evaluation units" staffed by a multidisciplinary team consisting of psychologists, learning disability specialists, social workers, psychiatrists, and guidance counselors. To avoid racial and cultural bias these teams adapted a form of Mercer's (1973) concept of pluralistic assessment. Hence, a social worker obtained background information on a child's developmental history and family environment; an educational evaluator assessed academic strengths and weaknesses; a psychiatrist conducted an interview; a neurologist evaluated the child's neurological functioning when appropriate; and a psychologist administered a series of tests, the WISC, Bender-Gestalt, and such projective techniques as the TAT and Rorschach, to measure the child's intellectual and emotional functioning. In addition, some children were observed in a simulated learning environment called the diagnostic classroom. In a conscious attempt to prevent individual bias, the evaluation team met in conference to review all information and to reach consensus on classification and placement. Finally, a separate committee reviewed the evaluation unit's decision in a meeting to which parents were invited. Despite this procedure, the plaintiffs alleged that the school system placed excessive reliance on tests which they claimed were discriminatory.

The Court analyzed the plaintiffs' charges under a right to treatment theory, finding that right implied in the due process and equal protection clauses of the Constitution, 1964 Civil Rights Act, PL 94-142, and Sec. 504 of the 1973 Rehabilitation Act. It said, "Since proper evaluation is . . . central to acceptable special education, a program falling substantially below minimum established standards would constitute a violation of the right to treatment" (p. 1285). The court found that some of the defendants' assessment practices fell short of those standards.

The Court did not condemn the tests or examine claims of test bias, however. It found possible discrimination only insofar as most assessments were not performed by minority professionals. Rather, the court focused its attack primarily on the long waiting lists of children to be evaluated and placed, the lack of systematic annual review of students, and the transfer of students from one special education program to another without full diagnostic evaluations. It ruled that the school system violated requirements in PL 94-142 and Sec. 504 for triennial reevaluations to determine whether children should be retained in current placements or integrated into the mainstream of regular education.

Thus, as in the early school desegregation cases, the Court never scrutinized the psychometric soundness of the instruments used to assess these minority children for placement. For example, it did not inquire whether the Bender or Rorschach were valid for the purpose of classifying minority children as emotionally disturbed. Instead, it condemned omissions in established evaluation procedures and the shortage of personnel to carry them out. The ultimate result, for good or ill, was to mandate, rather than restrict, further testing.

What this and other cases teach us is that, in fulfilling the requirements for nondiscriminatory assessment, school psychologists and other education officials must pay as much attention to the process as the substance of assessment. Thus, the final subsection here focuses on one crucial procedural requirement: providing notice about an impending evaluation and obtaining consent for performing it.[5]

### 1. Notice and Consent

The procedures for providing notice of and obtaining consent for the evaluation of handicapped children, regardless of race or ethnicity, are found in PL 94-142. The pertinent provision is Sec. 121a.504 of those regulations:

> (a) **Notice.** Written notice . . . must be given to the parents of a handicapped child a reasonable time before the public agency:
>
> (1) Proposes to initiate or change the identificaton, evaluation, or educational placement of the child or the provision of a free appropriate public education to the child, or
>
> (2) Refuses to initiate or change (the above).
>
> (b) **Consent.** (1) Parental consent must be obtained before:
>
> (i) Conducting a preplacement evaluation; and
>
> (ii) Initial placement of a handicapped child in a program providing special education and related services.

---

*Task 3*

*In a brief paragraph, distinguish between* **notice** *and* **consent.**

---

The intent of Sec. 121a.504 is to increase parental participation in educational decision making. School systems must notify parents before they take certain actions and they must obtain consent before they engage in others. The most comprehensive and accurate evaluation performed without proper parental approval may be invalidated. Thus it is important to understand the difference between notice and consent.

Notice and consent are not equivalent. To notify is to *supply information* about impending actions. Consent requires *affirmative permission* before actions can be taken. This distinction raises three questions:

1. What is the legal definition of consent?
2. In what particular situations is notice sufficient; in which is consent required?
3. In obtaining notice or consent, what information does the law require school systems to give to parents?

Answers to the first two questions are provided in this section. The answer to the third is discussed in the last subsection.

#### a. Meaning of Consent

The concept of informed consent, although difficult to define, possesses three basic considerations: (a) knowledge, (b) voluntariness, and (c) capacity.

1. **Knowledge.** The person seeking consent must disclose sufficient information in a manner that can be understood by the person from whom the consent is sought. The school need not inform parents of every possible detail in the procedure for which consent is necessary. The regulations require only the communication "of all information relevant to the activity for which consent is sought" (Sec. 121a.500). What, then, is "relevant information"? Sec. 121a.505 of the regulations lists those items that school systems must disclose to parents prior to obtaining their consent (see following subsection, "2. Content of Notice and Consent Forms"). Full disclosure of every conceivable aspect of an evaluation is not required because it is impossible and not because it is ethically undesirable.

This restriction on the duty to disclose does not excuse school psychologists from making every good-faith attempt to inform parents of those items they must disclose. Language differences between the school psychologist and parents cannot be a barrier to communicating relevant information. Information must be imparted in the "native language" of the parents or in another mode of communication if the parents are unable to understand oral or written language of any sort. The term "native language" is defined, albeit vaguely, in the Elementary and Secondary Education Act of 1965 as "the language normally used by . . . individuals" of limited English-speaking ability.

2. **Voluntariness.** Communicating all necessary information in a comprehensible manner does not guarantee that consent will be voluntary. Consent must be obtained in the absence of coercion, duress, misrepresentation, or undue inducement. In short, the person giving consent must do so freely. From a philosophical perspective, it may be impossible to determine whether any decision is made freely, but, as a practical matter, consent is deemed to be voluntary if school officials do not use unacceptable influence or interference. Although psychologists may communicate the school's point of view and even attempt to influence decisions, the means by which the information is communicated should not destroy the parents' ability to weigh and consider that information.

3. **Capacity.** Persons must be legally competent to give consent. Children are considered incapable of making many legally binding decisions. Some adults, especially those in institutions, also may be judged incompetent. School personnel must be acutely sensitive to the civil and constitutional rights of persons they believe to be incapable of giving consent. Adults are considered legally incompetent only after they have been afforded a full hearing and an impartial factfinder (e.g., a judge) decides that they meet the legal test of incompetency. Psychologists should be very wary of questioning the competency of persons from whom they seek consent. When a court declares someone who is a parent to be incompetent, it may appoint a "guardian of the

person" to act as a proxy decision maker. The result is that the parent loses the right to make decisions for him/herself and the children.

Granted that capacity is a vital element in securing legally sufficient consent, school psychologists should assume that parents are competent. Language difficulties, mild intellectual impairments, problems in living, or refusal to consent is not enough to trigger an adjudication of incompetency.[8] On the other hand, consent from parents who clearly cannot comprehend any communication or who cannot respond in any manner to requests for consent does not meet legal standards. In such cases it would be appropriate for schools to initiate procedures for the appointment of a substitute decision maker who can represent meaningfully the interests of the parents and their children.

### b. When Consent Is Required

Consent is required only when the school seeks to conduct a *preplacement evaluation* and when it proposes the *initial placement* of children in special education programs. Because the latter is not the direct responsibility of the school psychologist, the discussion here is restricted to the first situation.

A preplacement evaluation is defined as a "full and individual evaluation of the child's educational needs" (Sec. 121a.531). Thus, large-scale screening of children to identify those who may be handicapped and need individually focused assessment fall outside this definition, although school systems must inform parents of impending screening. Classroom observation, which is used to assess teacher-child interactions or for screening purposes also falls outside the definition. The participation of children in large-scale screening and classroom observations is minimal and has no immediate or direct negative effects upon them.

When an assessor observes members of a group acting in public, there is, at best, an inconsequential invasion of privacy. However, when a particular child becomes the focus of an assessment whose effect or intent will be to recommend placement in a special education program, then parental consent must be secured for all procedures, including testing, interviewing, and observaton. Although, once the child is placed, regulations do not require consent for evaluations, merely notification to parents, consent should be obtained for any evaluation the school performs, except when the instruments are used to assess academic performance only (e.g., reading, writing, spelling skills). To insure compliance, school psychologists and other personnel responsible for the assessment should not proceed with an evaluation without evidence of either a signed consent form (or some evidence of oral consent) or a legal order authorizing the evaluation when parents refuse to consent.

### 2. Content of Notice and Consent Forms

Of the three components of informed consent, knowledge is the most important. If parents do not receive sufficient information to make a decision, the courts will look upon any consent as suspect. Fortunately, we do not have to guess what is important to tell parents. PL 94-142's regulations provide guidance on the information that must be communicated. Sec. 121a.505 states that parents are to be given descriptions of

- the action proposed or refused by the agency, an explanation of why the agency proposes or refuses to take the action, and a description of any options the agency considered and the reasons why those options were rejected;
- each evaluation procedure, test, record, or report the agency uses as a basis for the proposal or refusal; and
- any other factors which are relevant to the agency's proposal or refusal.

Thus, the knowledge component of consent can be satisfied if the school tells parents (a) that it proposes to assess their children with comprehensive, individual educational evaluations; (b) why it believes that the evaluation is necessary; and (c) the devices it proposes to use in the evaluation.

---

*Task 4*

*Choose three assessment techniques you typically use. Describe each in the language and form which you believe complies with Section 121a.505.*

---

The operative word in Sec. 121a.505 is "describe." To describe evaluation procedures is not merely to list them. Moreover, descriptions should not be framed in technical jargon. Sec. 121a.505(b)(1) requires that any notice or consent form be "written in language understandable to the general public." There follow some examples, on a continuum from minimally acceptable to preferable, of how the requirement of Sec. 121a.505 can be met. To do so we will use as examples an IQ scale, a reading test, and classroom observation.

1. **Minimally Acceptable**

   We plan to use the following procedures to evaluate your child who, we think, may be academically handicapped:

   (a) The *WISC-R*; an individually administered test of general intelligence.

   (b) The *WRAT* (Reading only); an individually administered reading test that measures word recognition skills.

   (c) Classroom Observation; we will observe your child in his/her regular classroom.

2. **Better**

   We plan to use the following procedures . . .

   (a) The *Wechsler Intelligence Scale for Children-Revised;* given to a child in a one-to-one situation. The test is designed to measure children's general intelligence by seeing how they do on a number of different kinds of smaller tests within the larger one.

   (b) The *Wide Range Achievement Test* (Reading only), an individually administered reading test that measures children's ability to recognize many unrelated words.

   (c) Classroom Observation; our school psychologist will observe your child in his/her regular classroom for 20 minutes at three separate times during the week of September 20, 1979.

3. **Preferable**

   We plan to use the following procedures . . .

   (a) The *Wechsler Intelligence Scale for Children-Revised.* This test will be given by the school psychologist in a one-to-one situation at which only the psychologist and your child will be present. The test is designed to measure a child's general intelligence. It has 11 smaller tests in it. One of these smaller tests tries to see how much children know about the world about them (such as the number of days in a week, the reason we celebrate the Fourth of July, etc.). Another test asks children to rearrange cartoon pictures in what the test publisher believes is the correct order. Another asks children to remember and repeat a string of separate numbers that will be read to them.

   (b) The *Wide Range Achievement Test* (Reading only). The test has three parts (reading, spelling, and arithmetic) but we will only give your child the reading portion. The reading test measures how well children can recognize words printed on a page. The words do not appear in a story but are printed in a list of unrelated words. The test does not find out how well children can understand a story, only how well they can read separate words.

(c) *Classroom Observation*. Our school psychologist will visit your child in his/her regular classroom during three reading sessions for about 20 minutes each time during the week of September 20, 1979. The psychologist will observe the class to see how well your child reads out loud from books, to see how well he/she listens to the teachers, and how much he/she may be bothering other children while they work. The psychologist will also look to see how effective the teacher is in teaching your child and in controlling the class.

### 3. Interim Summary

School psychologists must be willing to engage in *mutual* disclosure with parents if they are to act ethically, conform to the requirements of the law, and develop intimate, trusting, and honest relationships. The rules on notice and consent place administrative burdens on school systems, but they lead to an open and informative atmosphere that engenders parental cooperation, fewer challenges of tests as discriminatory, and fewer costly and time-consuming hearings.

## Conclusion

At least three benefits can be discerned from the increased involvement of courts and legislatures in the testing controversy generally and the issue of nondiscriminatory assessment particularly:

> It has made the profession, as well as society generally, more sensitive to racial and cultural differences and how apparently innocent and benign practices may perpetuate discrimination. It has sensitized psychologists and other mental health professionals to the fact that they will be held responsible for their conduct. Through the accountability mechanisms now inherent in the procedural protections afforded handicapped children and their parents, psychologists who work in school settings find that they cannot view themselves only as passive recipients on orders from their supervisors. To protect the rights of their clients, to safeguard their own integrity, and in the long run, to serve the asserted ends of their employers to effectively educate students, they must question their practices, their interpretations, and their ultimate recommendations. Finally, the attack on psychological testing has accelerated the search for alternative means of assessment so that what is said about children is a more valid, truer depiction of how they perceive themselves and how they function in all spheres of life. In that light, the intense and searching examination that psychological testing has received from the legal system should be viewed as both salutary and welcome. (Bersoff, 1979)

---

[1] The federal government also cannot deny to persons equal protection of the due process clause of the Fifth Amendment, under an interpretation by the Supreme Court.

---

[2] Of course, there were not data that validated the determinations of the clinicians based on those individual tests. The tests themselves were not identified in the court's opinion.

---

[3] In *P. v. Riles* (Phase II) (1979), the Court acknowledged that intelligence tests could not measure innate ability. "IQ tests, like other ability tests, essentially measure achievement" (p. 38).

---

[4] Similar assurances are required elsewhere. The Office for Civil Rights (D/HEW) has interpreted the 1964 Civil Rights Act to compel all school systems receiving federal financial assistance to develop procedures preventing disproportionate over- or underinclusion of children of any race, color, national origin, or sex in all special programs (see Oakland & Laosa, 1977).

Sex discrimination is not a predominant concern in these laws. However, in Title IX of the Elementary and Secondary Education Act and its regulations, Congress and D/HEW provide some restraints in this regard. For rules regarding sex-fair tests see Code of Federal Regulations. Vol. 45, Section 86.21(b)(2).

---

[5] For a comprehensive discussion of the procedural safeguards accorded parents and school systems under PL 94-142 see Bersoff (1978).

---

[6] The regulations for PL 94-142 provide mechanisms for contesting parents' refusal to consent. A full discussion of these procedures is beyond the scope of this module, but for ways in which the school may override a parental veto see Sec. 121a.504(c) and Bersoff (1978).

## THEORY INTO PRACTICE: SIMULATIONS

### Simulation One

Attend at least five case conferences in which EMR placement is at issue. They should be meetings in which you are not directly involved. Observe the decision making processes. Determine the bases for the placement decisions and then answer the following questions:

1. In each case, what was the primary determinant for placement?
2. In how many cases was IQ the primary technique relied on?
3. If IQ was not the primary determinant, what aspect of the assessment in each case "loomed as the most important consideration in making assignments to EMR classes"? (See discussion of *P. v. Riles.*)

### Simulation Two

This letter was written to a parent prior to an initial evaluation of a minority child suspected of being handicapped. Read the letter and then complete the exercises after it.

Dear Parent:

We believe your child, Donna, needs some special help. Federal and state law mandates that we offer that help. However, before we can provide this necessary remediation of your child's academic deficiencies we must first assess her. This we are planning to do. Therefore, our school psychologist wishes to perform a comprehensive, multifaceted assessment on your child. She will be given our usual battery of tests including the Stanford-Binet, the Bender-Gestalt, some projectives, and a reading test.

Because it is important that your child have this evaluation we want to accomplish it as soon as possible. Therefore, we plan to do it one week from today. If you refuse to have the evaluation done, please call me before then. If I don't hear from you, we will assume you approve. But you should know that without this testing we cannot provide the special services your child needs. I anticipate your cooperation.

Yours truly,

Notice T. Porquale,
Principal

1. There are at least nine major flaws in this letter that directly contravene legal requirements. List at least seven of them.

2. Draft a letter to this parent that more clearly conforms to the law as you now understand it. Although the major flaws are listed in the next subsection, you should try to write the letter before you look at the list.

### Major Flaws in Letter

1. The fundamental error is that the parents are not given the opportunity to consent, to affirm the permission for the school to perform this preplacement evaluation. This letter is more like an advertising solicitation and would be unlawful.

2. Too much technical jargon and other language that parents would find difficult to understand are used in the letter.

3. The use of the term "special help" does not communicate enough information. Parents must be told that the evaluation may lead to a proposal by the school to place a child in a special education program. There is not enough information in this letter to meet the knowledge requirement.

4. The tests are not described; they are merely listed. Even the list communicates very little. Again, the knowledge component is not satisfied.

5. The threat not to provide any special services creates undue influence on parents and may vitiate consent because it would no longer be genuinely voluntary. In fact, the statement is not true. Schools can assess children and provide special education even if parents refuse to consent, assuming that they use proper hearing procedures.

6. The parents are not given enough time to respond, which may run afoul of the voluntariness requirement.

7. The school psychologist will be doing all the evaluation. Evaluation must be performed by a multidisciplinary team.

8. There should be no "usual battery" of tests. Each child's evaluation should be tailored to the suspected disability.

9. It is poor form to say that the "law mandates" special help for the child. The school appears forced to provide this help rather than assuming the role of the child's active advocate.

## PRETEST KEY

**Cases-Principles**

1. iv. Hobson v. Hansen
   i. Larry P. v. Riles
   ii. Brown v. Board of Education
   v. Washington v. Davis
   iii. Griggs v. Duke Power Co.

**True-False Questions**

2. F
3. F
4. F
5. T
6. F
7. F
8. F
9. F
10. T

**Written Statements**

11. When school systems do not receive funds under P.L. 94-142 but do receive federal financial assistance. Where these children are suspected of being handicapped.

12. (a) The school proposes to assess their child with a comprehensive, individual educational evaluation.
    (b) Why the school believes the evaluation to be necessary.
    (c) What devices will be used in the evaluation.

13. Comprehensive evaluation. Evaluations for placement must be conducted by persons from education, medicine, and psychology who assess children "in all areas related to the suspected disability, including, where appropriate, health, vision, hearing, social and emotional status, general intelligence, academic performance, communicative status, and motor abilities."

**Multiple Choice**

14. [illegible]

15. e

16. e

17. d

18. e

## REFERENCES

Anastasi, A. *Psychological testing* (4th ed.). New York: MacMillan, 1976.

Bernal, E. A response to "Educational Uses of Tests with Disadvantaged Subjects." *American Psychologist*, 1975, *30*, 93-95.

Bersoff, D. N. Procedural safeguards under Pub. L. 94-142. In Bureau of Education for the Handicapped, DHEW (Ed.), *Due process: Developing criteria for the evaluation of due process procedural safeguards provisions*. Washington, D.C.: DHEW, 1976.

Bersoff, D. N. Regarding psychologists testily: Legal constraints on psychological assessment in the public schools. *Maryland Law Review*, 1979, *39*, 27-120.

Cleary, T. A., Humphreys, L. G., Kendrick, S. A., Wesman, A. Educational uses of tests with disadvantaged students. *American Psychologist*, 1975, *30*, 15-41.

Holmen, M. G., & Docter, R. F. *Educational and psychological testing*. New York: Russell Sage Foundation, 1972.

Jackson, G. D. On the report of the Ad Hoc Committee on Educational Tests with Disadvantaged Students: Another psychological view from the Association of Black Psychologists. *American Psychologist*, 1975, *30*, 88-92.

Jensen, A. R. How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, 1969, *39*, 1-123.

Mercer, J. *Labeling the mentally retarded*. Berkeley: UCLA Press, 1973.

Oakland, T., & Laosa, L. M. Professional, legislative, and judicial influences on psychoeducational assessment practices in schools. In T. Oakland (Ed.), *Psychological and educational assessment of minority children*. New York: Brunner-Mazel, 1977.

## TABLE OF CASES

Armstead v. Starkville Municipal School District, 461 F.2d 276 (5th Cir. 1972).

Brown v. Board of Education, 347 U.S. 483 (1954).

Griggs v. Duke Power Co., 401 U.S. 424 (1971).

Hobson v. Hansen, 269 F. Supp. 401 (D. D.C. 1967) *aff'd sub nom.*, Smuck v. Hobson, 408 F.2d 175 (D.C. Cir. 1969).

Lora v. Board of Education of City of New York, 456 F. Supp. 1211 (E.D.N.Y. 1978).

P. v. Riles, 343 F. Supp. 1306 (N.D. Cal. 1972) *aff'd* 502 F.2d 963 (9th Cir. 1974) (preliminary injunction); No. C-71-2270 RFP slip op. (Oct. 16, 1979) (decision on merits).

Washington v. Davis, 426 U.S. 229 (1976).

## RECOMMENDED READINGS

This bibliography is in two parts: The first lists a restricted number of significant annotated references that are particularly germane to the field; the second is more comprehensive and its contents can serve as a reading list of important materials.

### Annotated References

Bersoff, D. N. Regarding psychologists testily: Legal regulation of psychological assessment in the public schools. *Maryland Law Review*, 1979, *39*, 27-120.

A comprehensive historical and legal analysis of school testing litigation. It presents the legal theories underlying challenges to intellectual and personality tests, analyzes and criticizes Hobson and Riles, and places the issue of test bias within a legal framework. A revised version of this article will appear in B. Sales (Ed.), *Perspectives in law and psychology: Vol. III: Testing and evaluation*. New York: Plenum, 1980.

Hobbs, N. (Ed.). *Issues in the classification of children* (Vols. I & II). San Francisco: Jossey-Bass, 1975.

The classic and most respected book in the field. Readers should pay special attention to the following chapters:

J. Mercer. Psychological assessment and the rights of children.
R. Burt. Judicial action to aid the retarded.
D. L. Kirp, P. J. Kuriloff & W. G. Buss. Legal mandates and organizational change.
W. G. Buss, D. L. Kirp, & P. J. Kuriloff. Exploring procedural modes of special classification.

Cronbach, L. J. Five decades of public controversy over mental testing. *American Psychologist*, 1975, *30*, 1-14.

An excellent historical perspective with special attention to the role of psychologist as scientist advocate.

Kirp, D. L. Schools as sorters: The constitutional and policy implications of student classifications. *University of Pennsylvania Law Review*, 1973, *121*, 705-97.

An early but not outdated review of the legal issues in classifying students by a leading scholar in education law.

Novick, M. R., & Ellis, D. D. Equal opportunity in educational and employment selection. *American Psychologist*, 1977, *32*, 306-20.

An excellent overview of the issues by a psychologist, who is an expert in measurement, and a law professor who has been directly involved in testing litigation.

Oakland, T. (Ed.). *Psychological and educational assessment of minority children*. New York: Brunner-Mazel, 1977.

Another comprehensive text edited by a school psychologist who has done significant research in the area. An extensive appendix presents reprints of major legislation, regulations, and guidelines published by the federal government. Readers should pay special attention to the following chapters:

T. Oakland, & L. M. Laosa. Professional, legislative, and judicial influences on psychoeducational assessment practices in schools.
T. Oakland, & P. Matuszek. Using tests in nondiscriminatory assessment.

Reschly, D. J. Nonbiased assessment. In G. D. Phye & D. J. Reschly (Eds.). *School psychology:* Perspectives and issues. New York: Academic Press, 1979.

An overview of both the psychometric and legal issues written especially for school psychologists.

### Reading List

Many articles that are found in legal periodicals should be understandable to readers who have completed this module and have read the annotated references.

APA Task Force on Employment Testing of Minority Groups, Job Testing and the Disadvantaged. *American Psychologist*, 1969, *24*, 637-50.

Bersoff, D. N. Silk purses into sow's ears. The decline of psychological testing and a suggestion for its redemption. *American Psychologist*, 1973, *28*, 892-99.

Bersoff, D. N. & Miller, T. L. Ethical and legal issues in behavioral assessment. In D. Sabation & T. Miller (Eds.), *Describing learner characteristics in special education instruction*. New York: Grune & Stratton, 1979.

Burgdorf, M., & Burgdorf, R. A history of unequal treatment: The qualifications of handicapped persons as a "suspect class" under the equal protection clause. *Santa Clara Lawyer*, 1975, *15*, 855-910.

Cleary, T. A., Humphreys, L. G., Kendrick, S. A., & Wesman, A. Educational uses of tests with disadvantaged students. *American Psychologist*, 1975, *30*, 15-41.

Comment, Constitutional law—Limiting the use of standardized intelligence tests for ability grouping the public schools. *North Carolina Law Review*, 1973, *51*, 1964-73.

Comment, Equal protection and standardized testing, *Mississippi Law Journal*, 1973, *44*, 900-27.

Comment, Ability grouping in public schools: A threat to equal protection? *Connecticut Law Review*, 1968, *1*, 150-67.

Dent, H. E. Assessing black children for mainstream placement. In R. Jones (Ed.), *Mainstreaming and the minority child*. Reston, VA: Council for Exceptional Children, 1976.

Deutsch, M., Fishman, T. A., Kogan, L., North, R., & Whiteman, M. Guidelines for testing minority group children. *Journal of Social Issues* (Supplement), 1964, *22*, 129-45.

Flaugher, R. L. The many definitions of test bias. *American Psychologist*, 1978, *33*, 671-79.

Horowitz, H. W. Unseparate but unequal—The emerging fourteenth amendment issue in public school education. *UCLA Law Review*, 1966, *13*, 1147-72.

Jorgensen, C. IQ tests and their educational supporters. *Journal of Social Issues*, 1973, *29*, 33-40.

Kamin, L. Social and legal consequences of IQ tests as classification instruments: Some warnings from our past. *Journal of School Psychology*, 1975, *13*, 317-23.

Kamin, L. J. *The science and politics of IQ.* Hillsdale, N.J.: Laurence Erlbaum, 1974.

Kirp, D. L. Law, politics, and equal educational opportunity: The limits of judicial involvement. *Harvard Educational Review*, 1977, *47*, 117-37.

Levin, B., & Moise, P. School desegregation in the seventies and the use of social science evidence: An annotated guide. *Law and Contemporary Problems*, 1975, *39*, 40-133.

McClung, M. Competency testing: Potential for discrimination. *Clearinghouse Review*, 1977, *11*, 439-48.

Note, *Hobson v. Hansen:* Judicial supervision of the color-blind school board. *Harvard Law Review*, 1968, *81*, 1511-27.

Note, *Hobson v. Hansen:* The de facto limits of judicial power. *Stanford Law Review*, 1968, *20*, 1249-68.

Note, Legal implications of the use of standardized ability tests in employment and education. *Columbia Law Review*, 1968, *68*, 691-744 (see responsive comments by F. Ruch & P. Ash, *Columbia Law Review*, 1969, *69*, 608-18).

Note, Segregation of poor and minority children into classes for the mentally retarded by the use of IQ tests. *Michigan Law Review*, 1973, *71*, 1212-50.

Note, The legal implications of cultural bias in the intelligence testing of disadvantaged school children. *Georgetown Law Journal*, 1973, *61*,1027-66.

Note, Constitutional requirements for standardized ability tests used in education. *Vanderbilt Law Review*, 1973, *26*, 789-821.

Note, Constitutional law-right of privacy-school program designed to identify and provide corrective therapy for potential drug abusers held unconstitutional. *Fordham Urban Law Journal*, 1974, *2*, 599-610.

Note, Equal protection and intelligence classifications. *Stanford Law Review*, 1974, *26*, 647-72.

Note, The Education for All Handicapped Children Act of 1975. *New York University Journal of Law Reform*, 1976, *10*, 110-51.

Oakland, T., & Phillips, B. (Eds.) Assessing minority group children. *Journal of School Psychology* (special issue), 1973, *11*, 291-401.

Recent Cases, Constitutional law-right of privacy-personality test used by school to identify potential drug abusers without informed consent of parents violates students' and parents' right of privacy. *Vanderbilt Law Review*, 1974, *27*, 372-81.

Recent Developments, Education-school-instituted program to identify potential drug abusers right to privacy. *Journal of Family Law*, 1973, *13*, 636-38.

Samuda, R. J. Problems and issues in assessment of minority group children. In R. Jones (Ed.), *Mainstreaming and the minority child.* Reston, VA: Council for Exceptional Children, 1976.

Sattler, J. M. *Assessment of children's intelligence.* Philadelphia: Saunders, 1974 (rev. reprint, 1975).

Shea, T. E. An educational perspective of the legality of intelligence testing and ability grouping. *Journal of Law and Education*, 1977, *6*, 137-58.

Sorgen, M. Testing and tracking in public schools. *Hastings Law Journal*, 1973, *24*, 1129-90.

Teitelbaum, H., & Miller, R. T. Bilingual education: The legal mandates. *Harvard Educational Review*, 1977, *47*, 138-70.

Yudof, M. G. Equal educational opportunity and the courts. *Texas Law Review*, 1973, *51*, 411-504.

Ziskin, J. *Coping with psychological and psychiatric testimony.* Beverly Hills, CA: Law and Psychology Press, 1975.

# NONBIASED ASSESSMENT: SOCIOCULTURAL CONSIDERATIONS

Ronald W. Henderson
Division of Social Sciences and
Oakes College
*The University of California*
*Santa Cruz*

# Table of Contents

Introduction ........ 35

Objectives ........ 36

Pretest ........ 36

Stereotypes and Expectations ........ 38

  Cultural Diversity and Ethnic Stereotyping ........ 38

  Characteristics of Minority Children and Their Families ........ 38
    Family Characteristics ........ 39
    Cognitive Style and Sense Modality Preferences ........ 39
    Motivation: Achievement and Affiliation Motives ........ 40
    Motivation: Fatalism ........ 41
    Summary ........ 41

  Teach Expectations and Classroom Interactions ........ 41
    Teacher Expectations and Culturally Different Children ........ 41
    Learned Helplessness ........ 43
    Overcoming Helplessness ........ 43
    Summary ........ 44

Simulations ........ 46
  Discrepancies in Role Expectations ........ 46
  Sociocultural Variations in Motivation ........ 47
  Family Characteristics, Learned Helplessness, and Alientation ........ 48
  Attention, Task Persistence, Learned Helplessness, Teacher Expectancies, Expectancies, Social System Perspective ........ 49

Model Responses to Tasks ........ 49

Pretest Key ........ 51

References ........ 53

Recommended Readings ........ 55

## Introduction

Times have changed. Despite the many false starts, underfunded programs, faulty assumptions, and unfulfilled promises of the era of the Great Society, events since the mid-1960s have served to emphasize the pluralistic nature of American society and to stress the need for educators to consider the social and cultural characteristics of school children. What was recently almost unheard of—not so long ago, remember, Mexican American children were punished for using their native language at school—has become a cliche. Educators now must devise curricula, instructional methods, and pupil services that are in tune with children's cultural backgrounds.

The admonition is deceptively simple. Even the easy part, which requires the curriculum to reflect content familiar to students from non-middle-class, non-Anglo backgrounds, is accomplished rarely at a level beyond the introduction of holidays and national foods. The part that implies that educational processes should be governed by considerations of cultural background has developed even more slowly. The difficulty arises from the fact that little is known about the inter-relations of culturally determined student characteristics and instructional processes. In fact, there is considerable controversy over just what characteristics constitute the cultures of the various minority groups which are served by the schools in the United States. Thus, a number of perplexing questions and issues dot the boundaries between ethnic cultures and American schools.

It would be useful if school psychologists had access to complete ethnographic descriptions of the populations with which they work so that they could be sensitized to possible cultural explanations for the behaviors and learning problems encountered in the schools. Such a set of descriptions is well beyond the limits of this paper. In many cases accurate descriptions are not accessible from any single source. The alternative—listing characteristics out of context—would be counter-productive. We already have witnessed the outcomes of that approach. During the 1960s anthropologists and sociologists were called upon to provide descriptions that would help teachers to understand their charges. The characterizations teachers learned often were more stereotyped than the conceptions they had held previously. For example, teachers were taught that Mexican Americans have difficulty making it in an Anglo world because they operate with a present-time orientation (Kluckhohn & Strodbeck, 1961), or that people of meager means in many different locations suffer from the effects of being socialized in a "culture of poverty" (Lewis, 1961, 1966). The culture of poverty was purported to interfere with achievement strivings because people with such backgrounds were motivated to seek immediate gratification of their needs rather than to defer their needs for more significant returns at a future point in time as, it was assumed, middle-class people do.

This is not to say that the ethnic minorities in the schools do not have distinctive subcultures but, rather, that glib assumptions about cultural characteristics may lead to damaging stereotyping.

The approach taken in this section of the module is to explore how stereotyped conceptions of minority children and their families may influence the expectations of educators. We then examine how these expectations interact with children's actual classroom behavior and approaches to academic tasks to produce perceptions and behaviors in the students which are either adaptive or maladaptive in instructional and assessment settings. Child behaviors that seem to violate the norms for classroom culture are particularly relevant here. The psychologist has a crucial role in identifying discontinuities between home and classroom cultures.

## OBJECTIVES

Upon completion of this section of the module participants should be able:

1. To recognize or recall the following concepts, as evidenced by responses to multiple choice items, or appropriate use of the concepts in written responses and group discussion:
   a. culture
   b. society
   c. social status
   d. social role
   e. ascribed and assigned status/role
   f. nuclear family
   g. extended family
   h. matriarchial family
   i. patriarchal family
   j. modality preference
   k. cognitive style
   l. field dependence
   m. field independence
   n. stereotype
   o. locus of control
   p. learned helplessness
2. To use information on the sociocultural characteristics of children and families as the basis for possible alternative explanations of learning and adjustment problems.
3. To pick out inconsistencies in social science descriptions of the cultures of minority groups and to identify the influence of uncritical acceptance of these descriptions on the formation of group stereotypes.
4. To test their assumptions concerning the cultural characteristics of minority children through discussions with parents and other community members.
5. To describe relationships between teacher expectancies and differential student-teacher interaction patterns.
6. To describe cultural, motivational, and cognitive characteristics attributed to minority groups, and to identify the limitations of social science generalizations about these characteristics.
7. To describe the relation between locus of control and learned helplessness.
8. To list and briefly describe procedures which the research literature suggests may be effective in the alleviation of learned helplessness.
9. In a role-playing situation, to demonstrate skill in resolving differences in the role expectations of teachers and parents and in reaching a consensus on shared goals.

## PRETEST

The following items enable you to test your present knowledge of the information and concepts that are discussed in the content subsection. Answers are provided at the end of this section.

### A. Cognitive Knowledge

1. The term culture refers to
   a. Literature, the fine arts, and other higher forms of human creativity and expression.
   b. All the beliefs and behaviors, including the products of behaviors, that members of a group learn and pass on from generation to generation.
   c. The institutions, such as law and religion, through which the interpersonal relations of members of a society are organized.
   d. The uniquely human capabilities that are universally shared by homo sapiens.
2. Among American social scientists the term society refers to
   a. The aggregation of people who share a common set of institutions.
   b. The customary behaviors and institutions that are shared by a specifiable population.
   c. Patterns of conduct that govern interpersonal relations in a human population.
   d. The segment of people in a community who control the major portion of power and resources.
3. An ascribed status is
   a. An earned position in a social system.
   b. The behavior one is expected to display by virtue of attained position.
   c. A position assigned in a social system on the basis of non-modifiable attributes.
   d. The behavior one is expected to display by virtue of assigned responsibilities in a social system.
4. From a social system perspective, mental retardation and emotional disturbance are considered to be
   a. Achieved social statuses.
   b. Ascribed social statuses.
   c. Achieved social roles.
   d. Acquired social roles.
5. Which of the following terms does not fit with the other three?
   a. Locus of control.
   b. Experimental control.
   c. Internal control.
   d. External control.
6. Evaluate the assumption that the single-parent, matriarchal family that is predominant among American blacks is historically derived from conditions of slavery, in which family groups were broken up and children

were reared in mother-headed households.

7. Describe the typical family form among Mexican Americans and contrast this structure with the modal Anglo-American family.
8. Children in some cultural groups are socialized to be more influenced by affiliative than achievement motives. Describe the educational implications of such socialization.
9. How does the theme of fatalism in Latin American cultures influence the performance of school children from these groups?
10. Describe how the authoritarian, patriarchal structure of Mexican American families influences the academic motivation of their children.
11. Differentiation theory suggests that socialization practices influence the development of the field dependence/field independence dimension of cognitive style. How do Anglo-American and Mexican American children differ on this dimension, and what instructional adjustments have been recommended to provide for these differences?
12. Develop a brief scenario to depict how the reciprocal relationship between a minority student's behavior and capabilities and the environmental conditions in a classroom (including teacher behavior and tasks) may lead to a pattern of learned helplessness.
13. Describe the relationship between teacher expectancies and teacher-student interactions, and indicate how these patterns affect minority children.
14. Explain how the progressive increase in the variability of school achievement of minority children may be accounted for. Use the concepts of *precursor skills* and *discontinuities in behavioral norms* in your response.

**B. Practice**

Respond to each of the following questions on your own practices by placing a plus (+) beside the statements that represent things you usually do, and a minus (-) beside the statements that do not represent your typical practice. These items assume that your practice includes consultation services to teachers as well as diagnosis and assessment functions for children. If a statement does not represent practices that would be possible within the structure in which you work, place a zero (0) beside it.

_____ 1. I attempt to interpret the behavior of children referred to me in their cultural contexts by reading relevant social science and educational literature on the groups served by the schools in which I work.

_____ 2. When teachers have difficulty making material interesting or meaningful to children, I encourage them to include information on customs and notable people of the children's cultural background in the curriculum and classroom displays.

_____ 3. I encourage teachers to take their students' cultural backgrounds into account by dealing with contemporary social issues that impinge on the lives of the children and their families.

_____ 4. When a minority group child fails to respond correctly to an item on an assessment instrument, I consider the possibility that the observed performance may not represent the child's actual level of competence.

_____ 5. I use social science information on the cultures of groups I work with as a source of hypotheses to explain behavior and learning problems presented to me, but at the same time I try to avoid stereotyped expectations by regarding generalized cultural descriptions with a healthy degree of skepticism.

_____ 6. When a child fails to respond to a problem on an assessment instrument, I consider the possibility that I may be observing the effects of "learned helplessness," based on a history of failure experiences, rather than an expression of the child's actual level of competence.

_____ 7. When children are referred for poor academic performance, I help their teachers to develop instructional approaches that will result in successful learning experiences.

_____ 8. When a child cannot achieve an educational objective, I help the teacher to construct a task hierarchy to determine if the child possesses important prerequisite skills.

_____ 9. I try to avoid cultural stereotyping by checking my hypotheses on possible cultural causes of classroom learning and behavior with parents and other participants in the child's culture.

_____ 10. Parents are involved in instructional planning. We seek consensus on those educational objectives that are considered to be important in both home and school. Proposed procedures for attaining these objectives also are discussed to ensure that cultural norms are not violated.

_____ 11. I provide teachers with information on procedures to teach self-management skills to children as a means of overcoming learned helplessness.

To score the self-assessment of your practice, count all pluses and minuses. Divide that total into the number of plus signs you marked. This provides a ratio score that avoids penalizing you for those functions you cannot carry out in your own situation (i.e., the statements you marked 0).

There is no absolute passing criteria. If you scored relatively low by marking a high proportion of statements minus (-), you might consider how you could incorporate the practices suggested by those statements into your own work.

## Stereotypes and Expectations

### Cultural Diversity and Ethnic Stereotyping

---

*Task 1*

*(Model Responses to Tasks are located at the end of this section)*

*Before you read this section, list the items educators most often use to describe the characteristics of children from two minority groups with which you are familiar. Identify the similarities and differences in characterstics enumerated in the two lists. How do these characteristics differ from a list that would describe majority group children? Which characteristics in each list cluster together in a meaningful pattern that could be justified as aspects of an ethnic culture?*

*Compare your responses to the listed items with responses of members and nonmembers of the groups you have identified. How do you account for similarities and differences in your conclusions?*

---

A brief review of sociocultural concepts is in order before we turn to an examination of the ways in which ethnic stereotypes may influence teacher behaviors and student responses.

Often, the term "culture" is used glibly to describe and explain behavior, but the concept is extremely abstract. People who have attempted to operationalize it for purposes of cross-cultural psychological research (e.g., Holtzman, Diaz-Guerrero, & Swartz, 1975) have found the task to be challenging. Numerous definitions have been offered for the concept of culture (Kroeber & Kluckhohn, 1959), but a common idea is that culture is composed of *habitual patterns of behavior that are characteristic of a group of people. Those shared behavioral patterns are transmitted from one generation to the next through symbolic communication* (Kroeber & Kluckhohn, 1952).

With their traditional emphasis on symbolic communication as the means by which customary behavior is transmitted, most anthropological definitions have overlooked the important process of modeling and observational learning as a major means of acquiring culture by neophytes.

Cultures can be described at varying levels of inclusiveness. Thus, we may speak of the culture of an ethnic group; or of socio-economic strata in which certain patterns of behaviors, values, and preferences may be shared. But subcultures are part of a larger whole (Laosa & DeAvila, in press), and behavior within groups may vary more than behavior between groups (Henderson, 1980).

Among American anthropologists, the term "society" designates an aggregate of people who live together in an organized population. The focus of the construct "society" is the people whereas "culture" focuses on the behaviors and traditions the people share.

The members of each culture or subculture hold behavioral expectations for a number of different statuses in the social order. These expectations define the roles people play as participants in the society. A status is the name of a particular social position (e.g., school principal) and the role is defined by a shared understanding of the behaviors that are appropriate for that position.

Some statuses are *ascribed*. One does not "earn" them by the demonstration of specific behaviors, skills, or credentials. For example, the status "child" is ascribed. Other statuses are *acquired*, such as that of school psychologist. Mercer (1973) argued that certain statuses which often are regarded as personal characteristics actually are acquired statuses. *Mentally retarded* is one such status. When an individual is assigned a status, the role definitions and expectations for that status are supported by socialization agents through a variety of social influence procedures.

Roles do not exist in isolation. They are defined by *mutual* expectations operating between pairs of roles, such as parent-child, teacher-child, or parent-teacher. In order for the transactions in role relationships to be smooth, the expectations must be shared. Some problems arising between teachers and students or teachers and parents may stem from differential role expectations, especially in the case of families from ethnic subcultures or socio-economic groups that do not completely share the middle-class norms that govern school practices (Parra & Henderson, 1977; Winetsky, 1978).

---

*Task 2*

*In what way may "mental retardation" be considered an achieved status? What implications does this view hold for school psychologists?*

---

Achieved statuses are earned by meeting culturally defined criteria. The individual's behavior is labeled (status designation) on the basis of behavior that deviates from the norms of a given social context (the classroom), but the label then comes to be viewed as characteristic of the individual. The person then is expected to behave in accordance with the achieved role, and others (teachers) behave toward the person in accordance with a set of shared role expectations.

Ethnographers have described the distinctive patterns of belief, behavior, and products that are found among well-defined, relatively homogeneous groups of particular cultures. Under such circumstances social structures can be analyzed and the status-role relationships identified rather comprehensively. But the groups being served by the schools in the United States are not homogeneous and they do not display an unambiguous cultural configuration for the educators who must plan instructional programs and services. Although there are patterns of behavior that constitute somewhat distinctive subcultures within the United States, two points bear consideration: (a) There is an extensive range of behavior within any United States subculture (Blackwell, 1975; Henderson & Merritt, 1968; Laosa, 1979). Given the behavioral heterogeneity of ethnic and racial groups, it is important to consider both the diversity and the similarities which are found in any particular group. When diversity is ignored, stereotyping is the inevitable result. (b) The quality of social science research on minority group cultures recently has been questioned. Specific criticisms by revisionist historians and minority group social scientists are considered in association with specific issues in the following section.

### Characteristics of Minority Children and Their Families

What are the distinctive cultural characteristics of the minority children and their families to whom educational and therapeutic processes should be responsive? There is no dearth of literature describing the family life and other cultural aspects of the various minority groups to whom the schools have a responsibility, but the conclusions are far from unambiguous. A sample of the issues raised by this research is presented here as a source of hypotheses for explaining instructional problems found in working with minority children, and as a caution against unqualified generalizations.

*Task 3*

*Educators often postulate that problems for which children are referred to school psychologists are the result of family influences. This is particularly true for minority children. Identify some of the ways in which family characteristics are thought to influence children's problems in school learning and adjustment. Then suggest alternative hypotheses to explain these problems. Test your assumptions in discussions with parents from the target community.*

**Family Characteristics.** Children's academic excellence and their academic problems are attributed to characteristics in the family unit. The fact that a high proportion of gifted children are from Jewish and Oriental families is thought to result from the high positive value afforded intellectual activities and the role of scholar in both cultures (Kirk & Gallagher, 1979). On the other hand, family characteristics have been blamed for the educational failures that plague disproportionate numbers of children from Hispanic and black backgrounds.

For example, the black family has been characterized as unstable and matriarchal (Moynihan, 1967). Unfavorable comparisons of black families with the nuclear family that is assumed to be normative for white middle-class Americans have been consistent with a sociological assumption that the nuclear family, with its division of labor, is the family form best suited to meet the demands of modern industrialized societies (cf., Parsons, 1949).

Many social scientists assume that the mother-centered family represents the continuation of a pattern which was established during slavery (Pettigrew, 1964). Revisionist scholars have begun to question these conclusions, offering long-neglected data to support their alternative interpretations. Even during the hardships of slavery, many black slaves were able to develop stable, two-parent families, and many black households in both southern and northern cities during the late 1800s were two-parent families (English, 1974).

In brief, social science statements about black families and their history often have been generalizations from small, unrepresentative samples. Often, they perpetuate old assumptions without studying black families directly. Recent evidence from such documents as census records suggests that unstable families headed by females have not been the rule during the history of the black family in the United States. New studies have attempted to identify positive aspects of black families. Only about one-third of American black families conform to the stereotyped matriarchial single-parent form (English, 1974). Even for those that do seem to fit the stereotype, the entire social and political matrix of life for black people must be analyzed to understand the forces that shape their lives. The stereotyped description of a single-parent, matricentric family is likely to be more misleading than instructive for school psychologists who have been called upon to help structure more favorable educational opportunities for black children.

Something akin to a pathology model also has been employed to describe Mexican American and other minority families. But, unlike the stereotyped black family, which is characterized as matriarchal and unstable, the Mexican American family is depicted in the social science literature as a stable, patriarchal, extended unit (Montiel, 1973). The husband is depicted as an authority figure in the home who demands and receives unquestioning obedience from his wife and children. This supremacy in the home is seen as a compensation for the husband's second-class citizenship in the outside world. He makes all the financial decisions, disciplines the children, and represents the family in dealings with the community. The wife is expected to be chaste and unworldly. She puts her husband's desires before her own. This pattern is so widely accepted as the norm for the Mexican family that the Spanish word *machismo* has become the standard term for designating male dominance in American English (Hawkes & Taylor, 1975).

It is assumed that this form of family structure, with minor modifications, is dominant among Americans of Mexican heritage. The empirical data fail to support that generalization, however. Deviations from these presumed norms usually are attributed to acculturation, or to the increasing financial independence of women, but there is evidence to suggest that the patriarchal pattern may never have been the norm (Griffith, 1948; Grebler, Moore & Guzman, 1970, cited in Hawkes and Taylor, 1975; Woods, 1956). Moreover, data from the 1930s suggest that the Mexican American family in the United States was anything but stable at that time. Data based on Mexican American families in Los Angeles, California, and San Antonio, Texas, indicate that a very small percentage of the families were of the extended type; the form may never have been as prevalent as social scientists have suggested (Montiel, 1973). As for the authoritarian, patriarchal stereotype, Hawkes and Taylor (1975) investigated the prevalence of male dominance among Mexican American farm labor families and found that, by a wide margin, the most prevalent mode of decision making and action was egalitarianism.

The knowledge that common assumptions about the nature of the Mexican American family are faulty should call into question statements that purport to explain how their children's ability to do well in school is damaged. Paternal authoritarianism, strong family ties, and a present time orientation are presented as antithetical to traits such as achievement, independence, and deferred gratification. These traits are considered essential to mobility. The sharply defined division of labor assumed to function in the Mexican famiy is also seen as detrimental to mobility and advancement in a industrial society (Heller, 1966). The curious reasoning in this argument may be noted by remembering that, for the black family, exactly opposite traits (e.g., maternal authority, weak family ties, lack of a sharp division of labor) are held to be responsible for the educational difficulties of their children and youth.

*Task 4*

*Before you read the following subsection, jot down your answers to the questions in this exercise. When you have finished the section, go over your responses and see if you have changed any of your original ideas. For each question, give the justification for your answer.*

- *Is cognitive style a sociocultural characteristic that influences children's responses to instruction?*
- *Is cognitive style affected by differences in the socialization practices of parents from different subcultures?*
- *Should cognitive style or modality preferences influence the design of instruction?*
- *Can cognitive style and modality preferences be diagnosed?*
- *Can psychological consultants provide instructional prescriptions on the basis of information on children's modality preferences or cognitive styles?*

**Cognitive Style and Sense Modality Preferences.** Socialization practices are among the behavior patterns that may differ across ethnic groups; it would be surprising if variations in child-rearing practices were not associated with cognitive and behavioral differences among children who have diverse backgrounds. Some social scientists argue that as a result of different socialization experiences children develop different cognitive styles or preferences for one sense modality over another. Inasmuch as cognitive styles and modality preferences are assumed to influence how people view reality, it is suggested that an optimum match be made between children's cognitive styles and the instructional styles of their teachers.

One perspective on modality preferences suggest that people and groups can be classified as to their preferred orientation to either a visual or oral-aural mode of involvement with the world around them. Visually-oriented people are said to be more object oriented and more predisposed toward the literacy tradition. Those whose orientation is predominantly oral-aural are said to interact with the environment in a more personalized way. It is argued that the high rate of school failure among children from minority subcultures may stem from the fact that the school curriculum and methods are more congruent with the literacy tradition of middle class families than with the preferences of people whose socialization is embedded in a more oral-aural tradition (Lewis, nd).

The field dependence/field independence (FDI) dimension also has been postulated as an influence on the school achievement of minority children. It is asserted that children who are socialized in settings where conformance to authority is emphasized (such as the stereotyped Mexican American family) tend toward field dependence whereas those with more egalitarian upbringing are more field independent. By definition "field independent" people are better at solving problems which require the removal of an important element from the context in which it is presented. Expanded to encompass social situations, field independence allows a person to be free from various forms of social influence and to have a more fully developed sense of individual identity than field dependent people. Field independent people are thought to have more complex cognitive systems, that is, to be more differentiated (Kagan & Buriel, 1977). Increasing differentiation of structure and function is the hallmark of developmental progress in most theories of development (e.g., Piaget, Freud, Lewin).

Some researchers interpret existing FDI research to show that field dependent people have better social and interpersonal skills than people who are field independent. On the other hand, field independent people do better on tasks that require cognitive restructuring (Laosa, in press; Witkin & Goodenough, 1977; Witkin, Goodenough, & Oltman, 1977). Some researchers believe that neither style has an absolute advantage over the other. They agree that each pole of the dimension has an adaptive value for given tasks and situations (Laosa, in press). This view, however, has not dominated because most developmentalists assume that increased differentiation is the basis for the enhanced ability to deal with cognitive complexity. Thus, the designation of any group as field dependent is implicitly pejorative.

There is a fair amount of evidence that Mexican American children tend to score as more field dependent than do Anglo children on certain tasks; however, the research has not clearly established that this pattern results from authoritarian socialization. Neither has it been clearly demonstrated that a causal relationship exists between a field dependent cognitive style and interpersonal skills. What then are the educational implications of the claim that Mexican American children are more field dependent than Anglo American children?

On the one hand, it has been advocated that schools train children to be more field independent because this style is more compatible with the analytic requirements of school tasks than the field dependent cognitive style (Witkin, 1967); others, e.g., Ramirez, 1973, have argued the need for a better match between the cognitive styles of students and teachers. In one study, Sanders and Scholz (1967) examined the hypothesis that field dependent and Mexican American children make better academic progress when they are paired with teachers with a matching cognitive style; contrary to expectations, he found that field dependent children with field independent teachers gained more than those with field dependent teachers.

Ramirez and Castaneda (1974) expanded on the FDI concepts, and they renamed the field dependent construct "field sensitive," a less pejorative term. They maintained that both styles must be represented in the classroom in order for education to be culturally democratic. Long-term positive effects in mathematics and reading scores have been reported for a program based on the systematic elaboration of these basic assumptions (Kagan & Buriel, 1977), but results have not been fully published or replicated. Some additional support for the notion that consideration of children's cognitive style may enhance educational outcomes is provided in a recent study by Doebler and Eicke (1979).

Thus, evidence on the socialization antecedents of cognitive styles is mixed, and the outcomes of attempts to take cognitive style into consideration in instructional practices are still inconclusive. It would be inappropriate to encourage educators to adopt stereotyped expectations that children from a particular background will have difficulty with certain kinds of school tasks because of a particular cognitive style that has been associated with group membership. On the other hand, when children have difficulty responding to particular kinds of academic expectations, it would be appropriate to explore the hypothesis that the difficulty may be exacerbated by a cognitive style that is not suited to the requirements of the task.

---

*Task 5*

*Assume that a teacher asks you for tips on organizing a more effective instructional program for children who are more motivated by affiliation need than by achievement need. What suggestions could you offer?*

---

**Motivation: Achievement and Affiliation Motives.** Motivation is too broad and complex a topic to be treated in detail here. Just two selected issues are touched on to illustrate the misleading assumptions that often are made about the academic motivation of minority children.

The first is that children must be activated by *achievement motivation* if they are to do well in school. Achievement motivation requires learners to be guided by internal standards of excellence. Because some Hispanic and Pacific Island groups are thought to be motivated more by the *need for affiliation* than the *need for achievement*, they are expected to do relatively poorly in schoolwork. Thus, for example, the affiliation motive, which is reported as characteristic of Hawaiian children, is assumed to account for their poor record of achievement. They are assumed to give higher priority to helping others, seeking good fellowship, or honoring personal commitments than to seeking personal gain; therefore, it is reasoned, they sacrifice individual scholastic achievement. Contrary to this expectation, a motivational pattern based on affiliation has proved not to be a negative factor in the school achievement of Hawaiians (Gallimore, 1974). Inasmuch as these children tend to be responsive to peers and are likely to work toward group goals, educators may be well advised to organize instructional activities that emphasize cooperative rather than competitive, norm-referenced learning tasks. This suggestion should be treated as an hypothesis to be examined in practice with a given group of children rather than as a prescription. Given the variability that exists in identifiable groups, a cooperative predisposition or affiliation need should not be assumed. Nor should the efficacy of any particular mode of structuring classroom goals be taken as a panacea.

---

*Task 6*

*A teacher refers a Mexican American child to you as an "underachiever" in the belief that the child lacks motivation to do schoolwork because "his people" have a fatalistic attitude that prevents them from exerting effort to overcome difficulties. Can you think of an alternative hypothesis to explain the "lack of effort" described by the teacher? Record your opposing hypotheses and your rationale for them now. Then read the following material on fatalism and the information in the following subsection, "Teacher Expectations and Classroom Interactions."*

*Does this information change or confirm your hunches?*

---

**Motivation: Fatalism.** Educators have been taught that an attitude of *fatalism* hampers the educational, social, and economic advancement of Mexican Americans (Heller, 1966; Madsen, 1964; Paz, 1961). It has been suggested that Mexican Americans view good or bad fortune as the work of fate, and that this fatalism leads to resignation, which Anglo observers interpret as lack of drive or determination. Some social scientists (e.g., Heller, 1966; Madsen, 1964) have suggested that whereas Anglos try to overcome the misfortunes that befall them, Mexican Americans accept them.

If this suggestion were true then it might well explain some patterns of behavior which are displayed by Hispanic children in school. But careful attention to available data suggests that the generalization is overdrawn. One quantitative study (Farris & Glenn, 1976) found that when level of education was controlled, there were no differences between Anglos and Mexican Americans on the dimension of fatalism. More viable explanations than a generalized attitude of fatalism may explain why some Hispanic or other minority children fail to respond to failure with increased levels of effort. (See following subsection, "Teacher Expectations and Classroom Interactions.")

**Summary.** In the population of children served by the schools there are numerous subcultural groups whose behaviors differ in many ways from the norms of middle-class America. There are problems in identifying what behavior patterns are characteristic of the culture of a particular group because the patterns often co-vary with other factors, such as socio-economic status or rural vs. urban life style.

Although an understanding of the cultural background of children with whom we work is important, blanket descriptions sometimes are more harmful than helpful because the social science research base for the cultural description of ethnic groups is suspect in several aspects. The foregoing review illustrates the need for cautious interpretation.

Many descriptions start with the acceptance of dated assumptions, without the benefit of first-hand study of the communities in question. Any particular behavior is likely to have meaning only in the context of a total pattern of customary behavior. Fragmented postulates, taken out of context, often lead to stereotyped expectancies. This is not to say that useful descriptions are not available, but much of the existing literature is adequate only to suggest hypotheses. When black children speak out during a teacher's presentation rather than remaining silent or politely raising their hands for recognition, an educator familiar with the culture of the black community may consider their response to indicate interest rather than disrespect. Similarly, when a child avoids eye contact with adults and looks down when he/she is spoken to by a teacher, an instructor of Anglo American background may interpret the child's behavior as an indication of sneakiness; but if the child is a Papago native American, a teacher familiar with their culture should recognize the behavior as a way of showing respect for elders.

### Teacher Expectations and Classroom Interactions

Children vary substantially in personal characteristics and capabilities when they first come to school. In addition to this individual variation, the normative behavior displayed by groups of children from diverse cultural backgrounds may differ from the norms of traditional school culture. The black child speaking out in class or the Papago child avoiding eye contact with teachers are examples of deviations from the role behaviors teachers are likely to expect of students.

Differences between the classroom cultural norms of teachers and the cultural norms of children from certain ethnic subcultures or lower socio-economic status are likely to result in conflict or in differential treatment of children, depending on how closely the behavior approximates the norms of school culture. Proponents of an ecological perspective suggest that prohibitions that have little to do with actual instructional effectiveness often become the source of difficulty between the *culture bearer* and *culture violators* (Rhodes, 1967), that is, between teacher and students. Furthermore, there is evidence that teachers tend to hold lower expectations of academic success for students who violate school norms by expressing disinterest or inattention than for students from the middle-class majority culture. These expectations often color teacher/student interactions.

**Teacher Expectations and Culturally Different Children.** Several years ago Rosenthal and Jacobsen (1968) created a stir in educational circles with their book *Pygmalion in the Classroom*. These investigators examined the self-fulfilling prophecy hypothesis that teachers adjust their instructional efforts to match their beliefs in the ability of individual children and the children's learning then conforms to the teachers' expectations. The two investigators attempted to alter teacher expectancies experimentally by providing false information on certain children, that is, that they had been identified as having hidden potentials. The investigators reported increases in the intelligence test performance of the falsely identified high-potential children. Unfortunately, the study was so badly flawed that the authors' conclusions could not be justified (Elashoff & Snow, 1971). Subsequent research on the self-fulfilling prophecy has often yielded results inconsistent with those reported by Rosenthal and Jacobsen. Nevertheless, their report was readily accepted by many educators and civil rights activists because it offered an attractive alternative to the existing explanations of school failure in which the blame was placed on minority children themselves, their families, and their genetic inheritance; however, other educators dismissed the entire hypothesis as implausible.

It is undoubtedly naive to assume, as Rosenthal and Jacobsen did, that the simple manipulation of information provided to teachers could produce a speedy influence on so general a measure as IQ. On the other hand, there is good reason to believe that reciprocal influences in the classroom may produce cumulative failure and "behavior problems" among many culturally diverse children. The brief review that follows provides evidence that teachers form differential expectations of the achievement and behavior of the children they teach, and that those expectations influence teachers' interactions with students. The question of what variables influence teacher expectations and how children's learning is affected also is explored.

---

### *Task 7*

*Before you read the next section, reflect on your conversations with teachers, discussions in the teacher lounge, and other situations in which educators discuss the behavior and performance of students. List the categories of student characteristics which you think may be determinants of teacher expectations.*

---

A substantial body of research (Adams, 1978; Brophy & Good, 1974; Laosa, 1978; Lockheed, 1977) shows that teacher expectations are associated with children's personal characteristics. These expectations apparently are based on stereotyped conceptions of various behaviors and, at times, of group membership. Even physical attractiveness influences expectations (Adams, 1978) so that teachers judge unattractive children to be less intelligent, to show poorer academic promise, and to be less well-behaved in the classroom than attractive children. Inasmuch as judgments of attractiveness are determined by cultural standards, children who display strikingly different physical characteristics may be subject to especially negative expectations. Teachers tend to judge the potential of unattractive children to be lower than that of attractive children for academic and social development; also, teachers appear to be more willing to recommend unattractive children for special class placement (Ross & Salvia, 1975).

Although some researchers assume that teacher expectations influence achievement, others argue that just the reverse is true, that is, that the student achievement determines teacher expectancies (Dusek & O'Connell, 1973; Dusek & Wheeler, 1974; Williams, 1976). Other evidence suggests that differential teacher expectations and behaviors are not influenced by their perceptions of student ability or achievement per se. Rather, teachers may be responding to student behaviors that are interpreted as reflecting academic motivation. Luce and Hoge (1978) found that when fourth-grade teachers interacted with students whom they judged to have low motivation for school work they were more procedural, more critical, and gave more behavioral warnings than when they worked with students whom they ranked higher in motivation. The student behaviors that seemed to serve as stimuli for the formation of teacher expectancies were task initiation and attention. As we have already seen, these behaviors are likely to be among those for which the behavior of many ethnic minority students differ from the norms of the school culture. Experimental research shows that attending and non-attending behavior have a marked effect on teacher behavior (Klein, 1971).

Willis and Brophy's (1974) work provides further insights into the ways in which the behaviors and attitudes of students and teachers may influence each other. These researchers found that the pupils to whom teachers felt *attachment* were seen as successful and compliant because teachers found their responses rewarding. Teachers expressed *concern* for those students who had difficulty with schoolwork but who were also compliant and reinforced teachers in interactions with them. Teachers responded by providing them with a good deal of remedial help.

Those students to whom teachers felt *indifferent* failed to respond in a way that teachers found rewarding. The nonrewarding responses led to a pattern in which teachers spent little time with the children, though they perceived that these students needed additional help. *Rejected* students not only failed to provide teachers with rewarding interpersonal contacts, but, also, they were credited with creating discipline problems and classroom disturbances. Teachers wanted to get rid of these students so they attributed low-ability traits to them which did not accurately reflect the children's demonstrated ability.

The cited investigations provide several kinds of evidence that teacher expectancies and attitudes may be more influenced by the observed behaviors of students than by the personal characteristics which are associated with group membership. Such evidence suggests that teacher expectancies may merely reflect previous experience with children displaying certain characteristics but it is instructive to note that teachers tend to express stereotyped expectations which are based on labels assigned to children. This is true even when the behavior observed in the labeled children is incompatible with the label (Foster & Ysseldyke, 1976; Gillung and Rucker, 1977).

For example, Foster and Ysseldyke (1976) asked teachers to list the behaviors they expected to be displayed by hypothetical children labeled emotionally disturbed, learning disabled, mentally retarded, or normal. As anticipated, the teachers held more negative expectancies for the children categorized with a deviance label than for normal children. Each teacher was then assigned to one of four groups and all groups viewed the same video tape of a normal fourth-grade boy engaged in a variety of test-taking activities and in free play. Each group was told the boy was a member of a different category. After viewing the tape, negative expectancies were expressed toward the children categorized with one of the deviancy labels, *even though* the deviancy label groups had observed behaviors that were inconsistent with the label. These results are relevant to consideration of possible teacher expectancy influences on minority and poor children because, traditionally, these children have been overrepresented in the assignment of special category labels.

---

### *Task 8*

*On the basis of your own experience, describe some of the ways in which you would expect instructional interactions to vary depending on differential teacher expectancies. Contrast teacher responses to high-expectancy students with responses to low-expectancy students.*

---

Differential teacher responses toward different categories of children have been well documented (Good & Brophy, 1974). Although clear that teachers may be reacting to differences in the achievement and motivational behaviors of children, the issue may be more complicated than that. For example, in one study of interactions between a white teacher and a small group of black and white nursery school girls, it was suggested that a lack of shared understanding of expectations and gestural meanings could account for the differences in the children's success in gaining the teacher's attention (Byers & Byers, 1972). However, merely providing children with teachers of matching ethnicity, which might seem to be the simple solution to this problem, is apparently not sufficient to change unequal treatment in the classroom (Byalick & Bersoff, 1974).

Thus, the research suggests that when children from cultural backgrounds diverge from school norms, they are likely to display characteristics that elicit negative expectations and patterns of teacher response which are designed more to manage behavior than to provide relevant skill and content guidance. Teacher expectations may be based, in part, on group stereotypes, but the teachers also may be responding to behaviors that deviate from the implicit norms which are reinforced in classrooms. As a function of apparent lower initial achievement and motivation, reinforced by teachers' stereotyped expectations which are based on factors such as race, ethnicity, physical appearance, or socioeconomic status, minority children may receive fewer skill and content-related communications from their teachers, resulting in poorer achievement, less task involvement, and diminished effort.

Bear in mind, however, that minority children are not a homogeneous group. They differ in their initial achievement behaviors and in the degree to which their behavior conforms to the norms of the school culture. Thus, an initial discontinuity between some children's entering repertoires of behavior may lead to initial failure, lack of support and the beginning of a series of reciprocal influences that result in cumulative discrepancies from the achievement and behavioral expectations of the schools. Children who fall into this pattern may come to feel helpless to influence their own lives in the academic context.

**Learned Helplessness.** Some individuals perceive themselves to be incapable of overcoming failure. They learn to feel helpless through experiences in situations in which they have no control over aversive events. The concept of "learned helplessness" has a common sense appeal to educators and psychologists because it seems to provide insight into the debilitating behavior of some school children who otherwise seem quite capable. In studies of this phenomenon, individuals convinced that an experimental task measured important intellectual abilities, were induced to perceive themselves as helpless when they fail on the task. The relevance of this learned helplessness concept to classroom settings is obvious.

The learned helplessness phenomenon has much in common with the idea of locus of control. According to the latter construct, individuals whose behavior is influenced by internal control believe that they have a substantial personal influence on the things that happen to them. They are likely to perceive the events that befall them as the result of their own ability or effort. Those people who fail to perceive a link between their effort and outside reinforcement are likely to perceive events as the result of luck or misfortune; they perceive the cause of events as *external*.

Some children seem to learn quickly, through their school experiences, that they are destined to do poorly in comparison with their classmates. They see no way for it to be otherwise. A disproportionate number of the children who develop perceptions of external locus of control in achievement situations are from minority or lower socioeconomic status backgrounds.

---

*Task 9*

*List some of the cultural characteristics of groups you work with which may make them susceptible to learning to feel that they are unable to cope with school tasks. Describe the responses of these children to academic tasks. What are the implications for testing?*

---

Both locus of control and learned helplessness exist along a continuum. Such designations as "helpless" or "external" merely are a convenience for referring to individuals whose responses tend to fall toward one pole or the other. Children who are characterized helpless tend to attribute their failures to a lack of ability rather than to the levels of their own efforts (Diener, Dweck & Rappucci, 1973). Even when helpless children initially are competent at a given task, once they have experienced failure at the same task they display maladaptive responses. Children who have learned to feel helpless in the face of difficulties tend to attribute failure to personal inability, and their performance on subsequent tasks is impaired. In contrast, non-helpless children are likely to attribute lack of success to insufficient personal effort and they display no deterioration in subsequent performances. In fact, the performance of children who attribute failure to lack of effort often shows improvement following failure, apparently owing to increased effort (Diener & Dweck, 1978).

Dweck's (1975) research on learned helplessness has clear educational implications. Because helpless children evidence little recognition of effort as a determinant of success or failure, in contrast to children who persevere even after failure, they fail to see their own responsibility for outcomes. Thus they are likely to see aversive situations as insurmountable and, following an unsuccessful effort, they may respond to subsequent tasks of the same sort by not trying harder or not sticking to the task. Dweck emphasized that an aversive event—in this case, failure on an intellectual task—is not in itself the cause of the helplessness phenomenon. The critical difference between helpless and persevering children appears to be in their perception of the relation between their own behavior and the failure outcome.

Histories of failure or success appear to play a major role in the development of internality-externality (Kifer, 1975; Weisz, 1979) and the more specific attributions of cause (e.g., effort vs. ability, task difficulty, or luck) that influence future expectations and task-oriented behavior. The nature of schooling in the United States is such that certain children may be predestined to experience heavy and repeated doses of failure from the first grade on. A disproportionate number of poor and minority children whose outside-of-school socialization is not highly congruent with the expectations of teachers in middle-class oriented schools are likely to fall within this group.

Thomas (1979) called attention to striking parallels between the features of the learned helplessness syndrome and the characteristics of children classified as learning disabled. Although the term "learning disabilities" refers to such a hodge-podge of symptoms that it is virtually impossible to identify the common characteristics of the children to whom the label is applied (Lilly, 1979), one common element seems to be an overlay of frustration and defeat. Learning disabled children often are portrayed as being convinced that they cannot learn, and much of the initial teaching effort with them is directed to motivating them to expend sufficient effort to achieve success (Thomas, 1979). Typically, these children are frustrated easily, are low in effort and persistence, and are unwilling to attempt even those tasks that are within the range of their ability. Their learning histories often are dominated by failure. The more consistent their history of failure, the more likely they will attribute failure to ability or lack of it (Frieze & Weiner, 1971, cited in Thomas, 1979).

---

*Task 10*

*List procedures a school psychologist in a consulting role could suggest to teachers for helping children to overcome feelings of helplessness*

---

**Overcoming Helplessness.** The provision of *purposes for learning* has been identified as an important mediator of the relation among locus of control, motivation, and performance. There also is evidence that the effects of purpose may vary, depending on whether learning objectives are set by a teacher or by the students themselves. Arlin and Whitley

(1978) tested the hypothesis that students would be more likely to accept personal responsibility for success or failure when they perceived a role in determining their own activities. The investigators anticipated that the perceptions of academic control and self-management opportunities would have an interactive influence on each other. The findings suggested that students who have been encouraged to manage their own learning are more likely than their peers from traditional classrooms in which the teacher set learning objectives to develop the willingness to accept personal responsibility for academic successes and failures. The results also suggested that opportunities for self-management may have more influence on perceptions of responsibility for failure than on perceptions of responsibility for success. The reason may be that in either type of classroom situation it may be easier for students to attribute success to themselves than failure, but rationalizations for failure may differ for the two types of classroom situations examined. When teachers determine the activities, failure may be attributable to bad luck or the fault of the teacher. Students who determine their own goals may find it more difficult to pin responsibility for their failures on external sources.

It seems unlikely that the effects of success and failure operate independently of the social situation in which such experiences occur. Classrooms constitute the major social context in which social comparisons of performance are made routinely. Classrooms that employ a competitive goal structure are especially likely to encourage social comparisons (Ames, Ames, & Felker, 1977; Henderson & Hennig, 1979).

Ames and her associates at Purdue set out to study how competitive and noncompetitive classroom settings influence children's beliefs about the causes of success and failure for themselves and others. Their findings confirmed the expectation that the effects of success and failure experiences depend upon the nature of the social setting in which the attempted performance takes place. Children's attributions, their judgments of who deserves what, and their satisfaction with the performance of self and others differed as a function of competitive and noncompetitive reward structures. Competition leads to self-derogation. Compared to children who failed in noncompetitive settings, those who failed under competitive conditions judged themselves to have less ability and to be less deserving of reward. They also experienced more negative effect than those who failed under noncompetitive circumstances. Ames and her associates suggested that "the consequences of failure are obviously negative, but the impact of failure in competitive conditions seems to be rather devastating to a child's self-perceptions" (Ames, et. al., 1977, p. 7).

Dweck (1975) designed a study to determine whether helpless children's perceptions of the relation between their behavior and failure outcomes could be altered with a form of attribution therapy. Beyond merely changing the perceptions of these children, Dweck was interested in determining if alteration in the children's failure attributions would result in the reduction of their maladaptive responses to failure. She compared the effects of Attribution Retraining to a Success Only procedure that was recommended by many behavior modifiers. Dweck anticipated that the Success Only intervention would improve the ability of helpless children to sustain their efforts despite failure because the treatment was expected to raise their expectations of success. Attribution retraining was expected to produce even greater improvement because it would provide a new interpretation for failure. Children subjected to this instruction would attribute failure to insufficient effort rather than to uncontrollable factors. Contrary to the investigator's expectation, consistent and sustained decreases in maladaptive reactions to failure were evidenced *only* by the Attribution Retraining group. Whereas these children were able to confront failure in a more adaptive manner, some of the children in the Success Only condition displayed *increased sensitivity to failure* after an exclusive diet of success experiences. All the subjects in the Attribution Retraining program showed increases in effort attributions, indicating that besides showing improved adaptation to failure in test situations, these children altered their attributions for failure with respect to mathematics in general.

The fact that children in the Success Only condition continued to display deterioration of performance following the intervention was interpreted by Dweck to suggest that the success only procedures, which many behaviorists advocate, may be shortsighted.

> An instructional program for children who have difficulty dealing with failure would do well not to skirt the issue by trying to ensure success or by glossing over failure. Instead it should include procedures for dealing with this problem directly (Dweck, 1975, p. 684).

Other research (e.g., Andrews & Debus, 1978) supports the contention that attribution retraining is effective in changing children's attributions and their resistance to extinction.

**Summary.** The evidence is clear that teachers entertain differential expectations for the achievement of children who vary in personal characteristics. The amount and quality of instructional interactions often differ along the same dimensions. The question of whether teacher expectancies are based on these personal characteristics per se or on achievement characteristics that happen to co-vary with these characteristics has not been answered with complete satisfaction. Some careful methodological work suggests that achievement is the determining factor. Even if that is the case, the results are the same. For example, if a relatively high proportion of poor and minority children enter school with achievement characteristics that elicit negative expectancies from their teachers, then it makes little difference whether the expectation was determined by achievement behavior or social class markers. The result is likely to be the same for the children concerned.

The information reviewed in this section suggests that a path model may provide a partial explanation for the common pattern of school achievement in which minority and poor children tend to fall progressively further behind their peers as they move through school. That is, the children's lack of prerequisite skills or lack of attention to instructional tasks may invite teacher responses that lead the children to fail and feel helpless in the academic situation. The result may be reduced effort, as evidenced in poor attention and task persistence, which further influences low teacher expectations and associated instructional responses. There is a considerable range of variation in subcultures in the degree to which the prerequisite skills and behavioral norms and expectations learned at home are congruent with the norms and expectations of the school culture. Thus, it is important to avoid stereotyped conceptions of the sociocultural characteristics of children from any given subculture, or from minority and lower socioeconomic backgrounds in general. Some findings (Laosa, in press) suggest that the most important factor in determining the congruence or incongruence of home and school socialization is the level of formal education attained by the mother.

Although the behavior and achievements of children may change, the research reviewed in this section indicates, that, once formed, teachers' impressions of student ability which are based on stereotyped expectancies are resistant to change, even when observable performance conflicts with expectations.

Some analyses suggest that achievement behavior is a major determinant of teacher expectancies but other work has demonstrated that differential teacher instructional behaviors may be associated more with judgments of students' motivations to do schoolwork than with teachers' estimates of ability or achievement in basic school subjects (Luce & Hoge, 1978). This finding is particularly interesting in connection with the knowledge that teachers are influenced markedly by the attending/nonattending behavior of students and with what is known about how failure influences children's subsequent approaches to tasks given their internal or external perceptions of causality. If helpless children respond to failure by declining to expend effort on subsequent trials, the negative perceptions of their ability may be compounded by the teacher's use of more controlling, critical, and externally determined influence.

Along similar lines, the work of Willis and Brophy (1974) suggests that when teachers and children share similar achievement characteristics, teacher behavior may vary as a function of student social behaviors, especially by the degree to which teachers experience the interactions as rewarding. Under these circumstances, students whose behavior styles differ from the middle-class norms of the classroom are likely to experience proportionally fewer supportive and content-relevant contacts with teachers. It would be no surprise if students from minority and poor family backgrounds were disproportionally represented among this group. The evidence suggests that discrimination of this sort is often unintentional, and that consultation that makes teachers aware of their differential interaction patterns may help some teachers to overcome the tendency in the findings summarized here (Good & Brophy, 1974).

The effects of failure on children's expectations and attribution of cause are meaningful only when they are considered in the social context. It means little to be unsuccessful at a task which can be accomplished by only a few individuals. But to do poorly on tasks that are defined as normative social expectations is likely to impact on children's perceptions of their own ability. School tasks are widely regarded as normative social expectations, and under the competitive goal structures and overt social comparisons that are implicit in norm-referenced assessment practices, failure is likely to be particularly damaging. Contrary to popular stereotypes, minority parents hold high academic aspirations for their children, but actual expectations often are curtailed by reality (Parra & Henderson, 1977). Children who experience failure in competitive settings, when compared to those who are unsuccessful in noncompetitive settings, are more likely to experience negative affect and to engage in self-derogation. Thus, their future strivings for achievement are likely to be discouraged. Competitive goal structures clearly highlight social comparisons and inhibit effort attributions.

Diminished effort is the natural consequence of attributions of outcome to inability. Children who learn to feel helpless in the face of difficulty attribute the difficulty to inability, which is detrimental to effort and persistence; their responses become maladaptive and performance deteriorates. Non-helpless children, in contrast, tend to attribute failure to insufficient effort, and their response is likely to be to exert more effort (Dweck, 1975). In fact, the critical difference between responses to failure by helpless and non-helpless (e.g., mastery oriented) children may be that the latter do not ordinarily make spontaneous causal attributions. Rather than seeking causes, they may pursue solutions through self-monitoring and self-instruction (Diener & Dweck, 1978).

A number of procedures which are designed to facilitate the adoption of internal attributions of cause, especially that of effort, have been tested with encouraging results. Because failure experiences seem to play a particularly important role in the development of attributions of inability and external causes, and in the learning of helplessness, the simplest solution may be to provide externally oriented and helpless children with a rich diet of success. The facts appear not to bear out this assumption, however. Although failure may be instrumental in the learning of helplessness, the removal of failure does not appear to constitute a sufficient or, perhaps, even constructive, condition to reverse the process. Attribution retraining results in sustained decreases in maladaptive reactions to failure whereas success-only experiences have been found to produce increased sensitivity to failure. This difference should be instructive for educators who attempt to help children develop feelings of efficacy in the context of schooling.

A variety of approaches, including attribution retraining, social reinforcement, and token systems, in combinations with social reinforcement, have demonstrated promise for effecting such changes. The performance of children in whom the locus of control is external seems to improve when clear purposes for tasks are communicated (Dollinger & Taub, 1977). Children who have a role in setting their own goals are likely to accept personal responsibility for success and failure to a greater degree than their peers for whom goals are set by their teachers (Arlin & Whitley, 1978).

A caution must be presented here: Merely changing children's causal attributions of failure from external to internal, or from inability to effort, is not likely to produce sustained desirable results unless instruction is arranged to provide opportunities for successful outcomes from the effort. In fact, to induce students to make effort attributions accompanied by energetic behavior is likely to have devastating results in the absence of opportunities for success. Hard work is a virtue of long standing in America, but Covington and Omelich (1979) have made a persuasive case that effort is a double-edged sword when it comes to school achievement. One of the few available defenses of a student facing academic difficulties is to avoid the implications of inability by refusing to try. Data collected by Covington and Omelich (1979) show that negative affect (shame) and attributions of inability were greater among college undergraduates following substantial effort than when they did little studying. These situations seem highly probable in cases in which students are presented with tasks for which they lack prerequisite skills; for them, failure may begin to set in early because traditional instruction so rarely provides for the careful identification and teaching of precursor skills and concepts that are required for the construction of behaviors that constitute instructional goals (Bandura, 1977; Bergan, 1977; Bergan & Parra, 1978; Bloom, 1976). Given such situations, it may be more adaptive for children to attribute failure to external influences than to their own inability.

## SIMULATION 1

## Discrepancies in Role Expectations

1. Type of Simulation: Role playing.
2. Purpose: To make school psychologists aware of the influence of conflicting role expectations on communication between home and school.
3. Material: Role descriptions for teacher, parent, and school psychologist.
4. Activity: A Mexican American parent meets with a teacher and a school psychologist to discuss the academic progress of a student who has been referred to the psychologist for "lack of motivation and academic failure." Each participant is provided with her/his own role description but no one has access to the role descriptions of the other actors. Following the session, a group of observers identifies positive features of the role-playing episode and makes suggestions for improved communication among the participants.

   (Note: The role specifications for this simulation are based, in part, on findings reported in Parra & Henderson, 1977. Trainers or workshop leaders may develop similar role-playing simulations based on descriptive materials pertaining to role perceptions in other groups.)

### Teacher Role

The teacher assumes that the child is uninterested in school work because the parents do not place a high value on education. He/She believes that children's intellectual potentials are dependent upon the kind of intellectual stimulation which children experience at home during the pre-school years. Inasmuch as the child in question is from a relatively poor Mexican American family, the teacher suspects that the student's home environment provided little intellectual stimulation during the early childhood years. The teacher suspects that the motivational problem may be partly a function of the child's having too limited a field of intellectual experience to relate to curricular activities. The teacher, although not sure that the effects of early environmental deficiencies can be entirely overcome, believes it would help if the parents were to expose the child to books, word games, and the like, and to provide strong reinforcement for any kind of academic effort the child might display.

In brief, the teacher feels that it is the role of both parents and teachers to foster intellectual development and motivation for academic achievement. The teacher imagines that these parens do not have very high aspirations for their children.

### Parent Role

The parents have been very concerned with the child's difficulty in school. They have been reluctant to initiate a conference because although they speak English, limited vocabulary, frequent confusion of pronouns, and difficulties with verb tenses and noun-verb agreements are a source of embarrassment in communications with "educated" people. (The parent attending the conference is aware of some of his/her specific problems with English because of recent participation in an adult education class.) Because he/she feels intimidated, his/her manner during the conference is one of diffidence.

The child's father works at strenuous manual labor and is periodically unemployed. The parents want something better for their children; they have dreams of the children going to college and entering professions. However, given their life circumstances and the child's present difficulties in school, the *expectation* that the parents' dreams may be realized differs from their *aspirations*.

Both parents value education but they do not see their role as that of a teacher of academic skills. In their view, teachers bear the primary responsibility for the intellectual and academic development of children. The primary responsibility of the home is to foster the child's social and emotional development. At the same time, the parents realize that children must function in a variety of settings, and they wish that the school would be more sensitive to their child's social and emotional needs. This youngster often comes home from school feeling defeated. The parents have almost stopped asking the child how school went today because the child has developed the feeling that he/she lacks the ability to do school work. The child admits to not trying anymore.

### Psychologist's Role

The psychologist is not well acquainted with the cultural background of the child in question but is sensitive to the possibility that cultural factors may be involved in the problem.

During the initial stages the psychologist takes the role of facilitator. He/She tries to keep communication open between the parent and teacher. Her/His overall aims for this initial conference are (a) to identify the problem from both the teacher's and parent's perspectives, (b) to determine the conditions contributing to the problem, from the perspectives of both parties, (c) to identify some goals that are of mutual importance to both parent and teacher, and (d) to set priorities for the goals so more detailed instructional planning can occur in subsequent conferences.

While the discussion unfolds, the psychologist decides to add a specific objective to the more general goals which have been identified. He/She tries to help the teacher and parent to develop some mutual role expectations to govern the interactions of the teacher, parent and student.

5. Suggestions for assessing the role playing:

   The group discussion of the role-playing episode should include consideration of the following questions:

   a. Which of the teacher's misperceptions (e.g., the expectation of low parental aspirations for the child) became apparent?
   b. What discrepancies in role expectations for the statuses of parent and teacher were revealed?
   c. What effective techniques did the school psychologist employ to achieve the conference goals and objectives?
   d. Did the teacher or psychologist behave in any way that might inhibit open discussion (e.g., criticisms of the child instead of objective discussion of skills or maladaptive behaviors, such as reluctance to try tasks, or facial expressions in reaction to nonstandard grammar)?
   e. Did the psychologist use appropriate procedures to help to move the discussion toward consensus (e.g., keeping the discussion focused, use of summary statements to show progress, etc.)?

f. In arriving at mutual goals for the child, did the psychologist make provision for both cognitive and social-emotional goals, and was it done with consideration for their interdependence?
g. Was the importance of providing the child with skills for self-management *as well as* success experiences mentioned?

## SIMULATION 2

## Sociocultural Variations in Motivation

1. Type of Simulation: Written communication
2. Purpose: To examine ways in which different types of motives may be used to promote academic effort.
3. Material: Memorandum from a building principal to the school psychologist.
4. Activity: Participants read the memorandum and
   (a) decide what additional information would be needed initially;
   (b) decide how to obtain needed information;
   (c) formulate an hypothesis to explain the problem behavior described in the memo; and
   (d) suggest an instructional alteration to be tested in the classroom.

**Memorandum**

TO: Mary Kahai, School Psychologist
FROM: Ken Kanaka, Principal, Bishop School
RE: Psychological Consultation

One of our 6th grade teachers, Ms. Jones, has been having trouble motivating students in her class. She is an experienced teacher and apparently was successful when she taught on the mainland, but this is her first year teaching in Hawaii. A number of children in her class are Hawaiian Americans. She has complained to me that the Hawaiian children seem to have very little achievement motivation and therefore are not likely to make much academic progress. She has tried to determine the current functioning level of each child and to individualize instruction on that basis. Thus she feels the work should not be too difficult, and she gives consistent social reinforcement for individual effort. Even so, the children cheat by helping one another whenever they think they can get away with it.

Ms. Jones has asked if she could try a token economy in her class as a means of increasing academic effort. I did not say absolutely no, but I did ask her to talk with you about some possible alternatives before she goes ahead with it. It seems to me that some of the problems with this class may come from her being a newcomer who is unfamiliar with the sociocultural characteristics of these children.

I would appreciate it if you would make an appointment to consult with Ms. Jones sometime soon. My main purpose in writing prior to your meeting with her is to give you a chance to think about my idea that Ms. Jones' lack of familiarity with the culture may be part of the problem.

5. Suggested Responses:
   (a) Information needed: You will probably need additional information on how work activities are structured in the classroom. It would be particularly interesting to know how goals and incentives are organized.
   (b) Obtaining the information: The needed information may best be obtained through informal classroom observation. Observations could be guided by the question, "Who benefits or receives reinforcement when children expend effort on tasks; the individual or the group?" Is the work of children compared on a norm-referenced basis?

   For purposes of this sample response, assume that you find that most work is reinforced on an individual basis, and that the performances of the children are compared with each other.
   (c) Hypothesis: Hawaiian American children will devote increased effort to academic tasks if the classroom goal structure is changed from the individualistic and competitive approaches that are now being used to a more cooperative strategy.
   (d) Rationale: Some research suggests that Hawaiian American children are very peer oriented. They often try not to stand out by doing better than their peers do, and they may achieve greater satisfaction from assisting a friend than from individual accomplishment. Therefore, try a form of instructional organization that structures opportunities for children to help each other and to work on cooperative learning tasks in which everyone in the group benefits. Changing the goal structure in a way that makes classroom activities more compatible with the cultural priorities of the children may be less intrusive to the instructional process than a token economy would be, and individually awarded tokens may be counterproductive, anyway, if the hypothesis is correct.
   (e) Consultant's suggestion: Try using the children's desire to support their peers and to subordinate themselves to the group by redefining what is meant by cheating. Make it possible for children to refine their own skills by helping each other. Set up some tasks so that each child can contribute at his/her level to group objectives. Reinforce the group for collective accomplishment. Avoid norm-referenced comparisons of the performances of individual children.

   Take observational recordings of on-task behavior for about a week before instituting the change. Continue to keep records after the change in procedures to see if the desired change in effort takes place. In this practical situation a reversal condition would probably not be tried. Most teachers, understandably, want to let well enough alone. (The school psychologist will have to devise a system for observational recording that will not be an unmanageable burden on the teacher.)

## SIMULATION 3

### Concepts: Family characteristics, learned helplessness, and alienation

1. Type of simulation: Case study.
2. Purpose: To examine options other than stereotyped explanations for behavior problems and poor academic performance.
3. Material: Case study with information on family background, school achievement, intelligence test performance, and teacher's and psychologist's interpretations of data.
4. Activity: Analyze the case study on the basis of what you know about sociocultural characteristics, patterns of teacher/student interaction, and learned helplessness. The analysis should suggest alternative interpretations and recommended actions.

**Case Study**

Andrew: Age, 13-1, Grade 7.

Andrew W. is a black 7th grader living in an inner city area. School records indicate that his mother is unmarried and that she supports her family by working as a hotel maid. Andrew has three sisters and a younger brother.

Andrew's teacher is concerned with his poor academic performance and disruptive behavior in the classroom. He seldom pays attention to instruction and finds as many ways as possible to avoid starting on an assigned task. He often dismisses an assignment by saying, "I can't do that." Even when he does start an assignment he seldom sticks with it for long. Instead, he moves about the room disturbing other students. When he does work on an assignment he rushes through it, putting down poorly thought-out answers and displaying minimum effort.

A standardized achievement test administered at the beginning of the year indicated that his grade-level performance was 5.2 in math and 4.1 in reading.

Academic achievement is not the only source of concern. The teacher thinks he shows signs of mild emotional or social maladjustment. This conclusion is based largely on Andrew's apparent inability to concentrate on school tasks and on his inability to control impulses in the classroom. For example, when he participates in a class discussion that interests him, he blurts out his ideas without awaiting his turn. On some occasions when the teacher "gets on his case" about that, he has responded by getting up and leaving school for the rest of the day. His attendance is marked by frequent absences. Anecdotal records in the file indicate that the teacher interprets these behaviors as signs of poor social adjustment and lack of respect.

Because Andrew is a poor student, the teacher regards him as an undesirable role model for his peers. She is distressed that the other students seem to look up to Andrew. This is of particular concern because Andrew has a great deal of influence with his peers. They gather around him and listen to his hip talk with what seems to be admiration. He can get most of the kids in his class to do about anything he wants them to.

Recently, his teacher referred him for testing. So far only one test, the WISC-R, has been administered. The verbal IQ was 84, the performance IQ was 100, and the full scale IQ, 90. The teacher wanted an MA score to get a notion of Andrew's developmental level. The psychologist reported a mental age of 11.6.

In a conference the teacher and school psychologist agreed that the problem was probably motivational because Andrew was not performing up to his potential, as indicated by the WISC-R. The teacher suggested that the lack of an achievement-oriented male role model in Andrew's home may be responsible for his lack of interest in academic work and for his failure to put forth the necessary effort to achieve. The psychologist agreed. Inasmuch as Andrew does not seem to respond to the teacher's attempts at positive social reinforcement, the psychologist finds it difficult to recommend an intervention that might overcome the effects of apparent deficiencies in the home environment. The recommended action was to meet with Andrew's mother and try to convince her that the boy should be assigned to the resource teacher on a "pull-out" basis.

5. Suggested Response:

The assessment of Andrew's basic problem as a motivational one is probably correct, but the explanation for it may not be.

Even though the WISC-R may be culturally biased, the fact that Andrew's academic performance is lower than one would predict on the basis of his IQ scores suggests that he is an underachiever. The WISC-R verbal IQ may be somewhat low as a result of the middle-class language tasks sampled by the test. This is suggested both by the higher performance score and by Andrew's apparent facility in the use of language to influence his peers.

Other explanations of Andrew's lack of academic motivation and disruptions of the class should have been explored. First, it is possible that Andrew has experienced a long history of failure on academic tasks, and he may be convinced that even with effort he cannot succeed. By finding ways to avoid trying or expending effort, he may escape the humiliation of attributing failure to himself. Given that teachers tend to respond more favorably to students who pay attention and try than to those who do not, the problem is perpetuated and Andrew's performance may show increasing deterioration.

When Andrew bursts into a discussion without awaiting his turn, it may be an indication of interest rather than lack of respect or social maladjustment. When this show of enthusiasm is rebuffed, Andrew is likely to see it as confirmation that his efforts always will be received negatively. Alienation is a natural response to such "no win" situations.

The hypothesis that the motivational problem derives from family circumstances is nonproductive. Even if one accepts the stereotyped explanation, educators cannot place an achievement-oriented male role model in the child's home and convert the family into the nuclear prototype so valued (despite its steady demise) by white, middle-class Americans.

The alternative hypothesis, that the observed pattern of behavior has been derived largely from the boy's experiences *in school*, lends itself more easily to instructional modification. Rather than remove Andrew's ability to influence his peers, make success possible for him by giving him a role in setting his own objectives and taking responsibility for monitoring his own progress. When Andrew finds that success is possible, the psychologist may try reattribution therapy techniques to get him to see the relation of effort to outcomes. An atmosphere in which cooperative rather than competitive goal structures predominate probably would facilitate the process, and provide opportunities in which Andrew can have a positive influence on his peers.

## SIMULATION 4

### Concepts: Attention, task persistence, learned helplessness, teacher expectancies, and social system perspective.

1. Type of simulation: Transcript of interchange.
2. Purpose: To identify and critique the central assumption of the interchange, with particular attention to implications of the social system perspective and the literature on learned helplessness and teacher expectancies.
3. Material: Transcript of an interchange between a teacher and a school psychologist.
4. Activity: Participants will read the transcript and (a) identify the main assumption regarding the problem behaviors discussed, and (b) suggest an alternative explanation that incorporates information on teacher expectancies, learned helplessness, and the social system perspective.

**Transcript**

The following transcript is a record of an interchange between a teacher and a school psychologist. The two are discussing a child for whom the teacher has requested psychological assessment. The child, George, is a black third grader.

Teacher (T): I am concerned with George. He isn't making much progress in this class and I think he may have a learning disability. He doesn't pay attention, and when he bothers to do his written work, he hurries through it without caring what answers he puts down. He just doesn't seem to take notice of any of the details of assigned exercises. It's bad enough that he doesn't pay attention to his own work, but he is continually out of his seat bothering other children. According to the teachers who had him before he has been this way since the very beginning.

School Psychologist (SP): What do you do when these things happen, when he doesn't pay attention, for example?

T: Well, I tell him to look at me and pay attention when I give directions, and to get it the first time because I don't intend to keep repeating the instruction.

SP: And what about his failure to attend to details in the written work you assign?

T: Well, I don't know exactly what it is. It's probably more than just one thing. As I said, he doesn't seem to notice details. But mostly I think he just doesn't care or isn't willing to put out the effort. He almost seems proud that he doesn't try. Then, of course, there is his hyperactivity. I suppose it's all part of the same problem. He just can't, or won't, attend. He's all over the room when he should be working. It seems as if I'm continually telling him to sit down and do his work and to stop disturbing the other children. Once I do get him back to work he's very distractible. Any little thing that happens will pull his attention away from his work. He just has no task persistence at all.

SP: Well, you could be right. There could be a learning disability. If he's having perceptual and attentional problems, for example, that could explain why he has trouble with details and tends to be so distractible. I will schedule him for assessment next week and maybe we can figure out what the problem is.

5. Suggested responses:

(a) George is being labeled as inattentive, distractible, and hyperactive. The tone of this conversation suggests that these characteristics are assumed to be qualities within the individual. The search for a learning disability diagnosis suggests a medical model interpretation of George's problem.

(b) If George has displayed this pattern of behavior for a long time, suggested by information from his former teacher, there is a good chance that teacher expectations regarding his potential academic performance have been low all along. Remarks of the present teacher suggest that much of her interaction with George is aimed at controlling his behavior rather than teaching him specific content or skills. Initial failure to conform to achievement and behavioral norms may have put George in a position of failing at school tasks, receiving responses that did little to help him to acquire necessary skills for successful performance, followed by further failures. If that were the case, he could feel incapable of doing the tasks presented at school. According to what is known about learned helplessness, diminished effort and lack of attention would be predicted in that situation.

If this scenario were true, George's behavior could be better understood from a social structure than from a medical model perspective. The labels being arrived at designate his status in the classroom social system. They are not traits that are intrinsic to his biological nature, or evidences of disease. The labels carry expectations that may influence George's continuing behavior and the behavior of others toward him in the school situation.

## MODEL RESPONSES TO TASKS

### Task 1

Responses to this exercise will differ depending on the groups chosen for comparison and your own knowledge of the groups in question. Past experience with this exercise suggests that the characteristics listed often reflect stereotyped views of ethnic and racial groups. Responses such as "lack of future time orientation," "uninterested in school learning," and "unable to delay gratification" are common with reference to several minority groups. Rarely can the characteristics listed be justified as broadly applicable generalizations about the "culture" of the groups named. As you will see in this section, conceptions that social scientists have offered to help educators to become more knowledgeable about the children they teach have sometimes contributed to cultural stereotyping.

### Task 2

Mental retardation is defined in relation to specific social situations. The behavior of an individual may be seen as deviant or subnormal in the school situation but the behavior of the same individual in another social setting may not be considered deviant at all. In the first of these situations developmental deviance may be considered as an acquired status. The person who occupies this status and is labeled as retarded plays that role, and other people relate to those who occupy the status of mentally retarded in accordance with the behavioral expectations for the role associated with that status. School psychologists should remember, and help other educators to understand, that from a social system perspective, mental retardation is a designation for a position in a social system rather than a quality of the individual. The same may be true for other statuses, such as emotionally disturbed.

### Task 3

The specific response to this exercise will depend on what groups are chosen for attention. In general, social scientists have used pathology models to describe the influence of a number of minority group families; the authoritarian, patriarchal form ascribed to Hispanics has been seen as providing socialization experiences that are detrimental to the development of the kind of motivation that is assumed to be necessary to academic success. Hispanic groups also have been described as fatalistic—a quality that is considered to inhibit the use of effort to overcome obstacles. It now appears that many of these conclusions were overgeneralizations from limited samples and do not represent conditions present in the majority of minority group families. Furthermore, long-range achievement and adjustment do not depend solely on the characteristics the child brings to school. Just as important is the response these characteristics elicit from educators. Over the long haul, the reciprocal relation is likely to be the determining factor. This suggests that educators should be aware of these interactions and make sure that children are not treated in a way that deprives them of skill and content-relevant interactions.

### Task 4

There appear to be group differences in preferences for modalities (e.g., aural-oral vs. literacy) and cognitive styles (e.g., field dependence/independence). However, these differences only represent differences between the averages of groups. There is a substantial range of individual variation within a given group on dimensions such as these. Thus, care should be exercised to avoid stereotyped assumptions that a given child will have a given style or preference on the basis of her/his group membership.

The presence of average differences between groups on both cognitive style and socialization dimensions does not necessarily mean that the socialization practices produced the preferences with which they are statistically associated. The causal relationship has not been firmly demonstrated for any given socialization pattern and style. Furthermore, ethnic group membership and socio-economic status are often confounded in studies that compare groups on dimensions such as these. Thus, one cannot assume that authoritarian child rearing, for example, leads to field dependence.

When there is reason to believe that children are unable to profit from a given type of curriculum or method of instruction, the suspicion may be a justifiable basis for experimenting with instructional adaptations aimed at providing materials and methods that are congruent with the child's approach to the processing of information.

Instructional adaptation is desirable but the utility of formal diagnosis to that process has not been clearly demonstrated. Cognitive style, for example, may be measured in different ways, and those different approaches do not consistently yield the same assessment. Moreover, the measures that have been used in most descriptive and laboratory research do not yield clear predictions of children's responses to different instructional approaches.

Informal behavior assessment of children's responses to given kinds of instruction is probably a better basis for adapting to the instructional needs of given children than "personality" measures. It is best to be sensitive to needs suggested by children's approaches to specific materials and instructional styles. Flexibility in style should be the aim. For example, if there really are group differences in modality preferences, children from a literacy tradition may be as much in need of aural-oral practice as children with aural-oral capabilities are in need of skills associated with literacy.

### Task 5

You might suggest that the teacher capitalize on the children's support for one another and on their concern for the group by organizing learning activities around a cooperative goal structure. As a school psychologist you also might help to design the means by which the teacher may assess the effects of this instructional adjustment on attitudes toward learning. Systematic observational recordings of time on task and task completion would be appropriate. Information on affective responses to the change would also be important.

### Task 6

It is not at all clear that Hispanic populations in the United States are any more fatalistic than other groups. Fatalism reported for these groups may reflect only greater feelings of futility by people of limited education or economic means, and people of Hispanic heritage are overrepresented in the lower economic segments of American society. Some evidence indicates that when level of education was controlled, Mexican Americans and Anglos did not differ on the dimension of fatalism.

Some children whose social behavior or preparation for formal school work differs from the norms of the classroom may seem fatalistic about their school work because a pattern of failure experiences sets in early. They may learn to feel helpless, or to believe that they lack the ability to succeed at academic work. Accordingly, they may try less because previous experience has taught them that they cannot overcome their academic difficulties with effort.

## Task 7

Research has shown that teachers hold different expectations for children who vary along a number of personal characteristics and behavioral dimensions. The following variables have been found to be associated with differential teacher expectations and behaviors:

student sex
student social class
student ethnicity
student English language proficiency
student physical attractiveness
student achievement
student academic motivation

It has not been clearly demonstrated that these characteristics themselves are the main determinants of teacher expectancies and responses to students. It seems reasonably clear that most teachers do not respond to children solely on the basis of their racial, ethnic, or socio-economic class identity. This section develops the argument that teachers are likely to respond to sociocultural characteristics that *vary within* subcultural groups.

## Task 8

Teacher's communications to students for whom they hold high expectations tend to be more supportive, more reinforcing, and more related to the skills and content of instruction than are their messages to students for whom their expectations are low. Conversely, teacher behavior toward low-expectancy students tends to be less supportive and more aimed at behavioral control and management, as compared with their interactions with children for whom they hold high expectations.

## Task 9

Students from culturally diverse backgrounds may lack the precursor skills and concepts necessary to succeed at school tasks. To the degree that the concepts learned in the family context differ from those assumed in the curriculum, children may be at risk to fall into a pattern of failure. Children who experience failure even when they try may come to feel helpless to overcome their difficulties with school learning. Consequently, effort may be reduced. Children who experience such feelings of helplessness often fail even at tasks that are within their capability. School psychologists involved in the assessment of children who have experienced repeated failure should take steps to make sure that they are testing capability, in so far as possible, rather than just performance. Careful attention to motivation and reinforcement of effort attributions may help. Even so, interpretations of test results should be tempered by the realization that children who have experienced repeated failures in school may be more capable than test results suggest.

## Task 10

A school psychologist serving in a consultation role could suggest the following approaches which have proved to help in alleviating feelings of learned helplessness:

1. Make sure the student is provided with a purpose for learning. "You will need it when you grow up" is a cop out.
2. Include the students in the decision-making process relating to the determination of goals and activities.
3. Teach self-management skills to support students' involvement in goal and activity selection.
4. Establish noncompetitive goal structures, in which norm-referenced comparisons are avoided.
5. Cue and reinforce effort attributions.
6. Provide reattribution training.

## KEY TO PRETEST

The code following the answers indicates the major concept/subconcept where the materials are located in the discussion section.

1. b. Culture is a broad concept encompassing all the learned, shared, and transmitted behaviors characteristic of a group of people. Culture can refer to the characteristic habits of a large group of people, such as those who are members of the society of the United States of America, or to more restricted groups, such as Mexican Americans or Sioux Indians. (Cultural diversity and stereotyping/culture)

2. a. As the concept is used by American anthropologists and sociologists, the term "society" refers to the collection of individuals who live together in an organized population. The focus is on the people rather than their behaviors. Society is distinguished from culture in that culture focuses on customary behaviors and products of behavior shared among people in a given society. (Cultural diversity and stereotyping/society)

3. c. A social status is a position in a social system. Available statuses may be named (e.g., child, school psychologist), and each individual occupies several statuses simultaneously. The behaviors that are expected of an individual occupying a given status constitute the role for that status. Ascribed statuses are based on characteristics that are not subject to purposeful modification. (Social status/social role)

4. d. From the social system perspective, children acquire various statuses on the basis of their behavior in one or more of the social systems in which they participate. These statuses include designations such as retarded student, gifted student, and emotionally disturbed student. Mental retardation and emotional disturbance are the designations for the roles (expected behaviors) associated with the social statuses of mentally retarded and

emotionally disturbed students. (Social status/social role)

5. b. Locus of control refers to an individual's expectations whether she/he has control over his/her circumstances, or whether such matters are beyond his/her personal control. The first of these conditions is referred to as internal control, the second, as external control. The term "experimental control" has nothing to do with individual locus of control perception. (Cultural diversity and stereotyping/locus of control)

6. Most social science literature has portrayed the black family as a single-parent family dominated by a mother. This family form has been identified as the source of various social problems. The claim has been made that this form is derived from conditions of slavery which broke up family units. Revisionist scholars have presented neglected data indicating that even in slave times, a sizeable proportion of blacks managed to maintain two-parent families, and the stereotyped single-parent matriarchal family does not represent the majority of black families in either the past or the present. (Cultural diversity and stereotyping/family characteristics)

7. Traditionally, the social science literature has presented the Mexican American family as a father-dominated authoritarian structure. The father's word was law and his wife and children were expected to behave with unquestioning obedience. The wife's place was in the home, and the father represented the family in matters outside the household. This family type usually is presented as though it were present with only minor modification in Mexican American and Chicano households. The Mexican family and the transplanted version of it are described as extended family units, with households including grandparents, married offspring and their spouses and children, and sometimes other relatives.

   This family form is usually contrasted with the nuclear family that is considered typical among middle-class Anglo Americans. A nuclear family consists of only a husband and wife and their own biological offspring.

   Mexican American scholars recently have challenged this characterization with data suggesting that even in Mexico the extended family is not as common as has been suggested, and that among Mexican American families, egalitarian values are dominant over authoritarian, patriarchal values. (Cultural diversity and stereotyping/family characteristics)

8. A widespread belief in psychology is that achievement motivation is essential to academic and economic advancement. Affiliation motivation, on the other hand, has been considered to be detrimental to the kind of individualistic achievement that is necessary for academic success. Those who are more motivated by affiliation needs than by individualistic achievement strivings are likely to put concern for interpersonal relationships and obligations over economic personal gain or a competitive advantage in scholarship. Some recent research with Hawaiian groups suggests that affiliation motivation need not interfere with academic achievement. Positive correlations between affiliaton motivation and achievement have been reported. Affiliation motivation and achievement motivation may be alternative systems, each of which may have advantages within the value context of given cultures. (Cultural diversity and stereotyping/maturation)

9. Much of the literature on Latin American culture has stressed the fatalistic outlook of people from these groups. When an individual fails, it is seen as the work of fate rather than a circumstance to be overcome. Some research suggests that the fatalistic orientation described for Puerto Ricans and Mexicans may not be applicable to Hispanic families in the United States. (Cultural diversity and stereotyping/family characteristics)

10. The popular stereotype of the Mexican American family as a structure in which children cannot question authority goes on to suggest that these rigid socialization patterns produce uncreative children with field-dependent cognitive styles. Mexican American child-rearing practices also are purported to inhibit initiative and individualistic achievement patterns. Revisionist scholars have condemned these overdrawn generalizations. Research with Mexican American families has demonstrated a substantial degree of heterogeneity. (Cultural diversity and stereotyping/socialization)

11. Research on cultural differences in cognitive style has shown with some consistency that Mexican American children tend to be field dependent whereas Anglo children tend to be more field independent. In each group, however, the range of individual differences is substantial, and some research has failed to show such differences.

    Some investigators advocate that children be taught by teachers whose cognitive style matches their own. The research on this approach is inconsistent in its results. Others advocate that teachers be made sensitive to both cognitive styles and that instruction be organized to help children to function "bicognitively." This approach awaits thorough evaluation and replication. (Cultural diversity and stereotyping/cognitive style)

12. Because of incongruities between the cultures of home and school, a minority child may lack prerequisite skills that many teachers take for granted and never teach directly. Thus, the child's chances for initial failure may be greater than those for a middle-class child. The greater the discrepancy between the cultures of school and home, the greater is the possibility of initial failure.

    Aside from the possibilities for early failure on standard school tasks, teachers often hold expectations of poor classroom behavior and less potential for academic progress for children from minority and lower socio-economic groups. The combination of the child's difficulty with initial school tasks and the teacher's expectations may influence the teacher to spend a disproportionate number of contacts with this child managing his/her behavior rather than communicating curriculum relevant content and skills. As a result, the child may fall further and further behind, and a pattern of failure may set in. If the child has been expending effort on the school tasks, he/she may come to believe that failure is the result of a lack of ability to do academic work. A maladaptive pattern whereby tasks are not initiated and effort is withheld may develop as the child comes to feel helpless to overcome these difficulties.

    The perception of helplessness may be limited to academic tasks. A black child with a good jive walk and facility with hip talk may be a leader in the peer group even if not the shining star of things teachers hold dear.

(Teacher expectations and student performance/learned helplessness)

13. Teachers often hold expectancies that minority children will fail to make good academic progress, and that their behavior is likely to be more disruptive than that of non-minority students. These expectancies may be influenced by a variety of circumstances, including special labels applied to children, and actual observation of behavior and achievement. Whatever the determinant of the expectation, the teacher is likely to spend less time in supportive interactions with these students, and more time in efforts to dominate and control social behavior. Communications are likely to include less information about skills and information the child needs to progress academically, and more behavior management concerns. Minority children who fall into this pattern of interactions with their teachers are deprived of full access to the curriculum. (Teacher expectations and student performance/differential student/teacher interaction patterns)

14. There is considerable variability in children's behavior when they enter school. If teachers and psychologists do not take special pains to ensure that children have the necessary skills to accomplish the tasks presented to them, they have a slim chance of accomplishing the instructional objectives. Each objective may be prerequisite to the others, so the child in this condition slips progressively further behind those children who possess the prerequisite skills.

    Children whose behavior violates the cultural expectation of teachers also are likely to be at an initial disadvantage in the classroom, and the effects of the disadvantage may be cumulative. Culture-violating behavior becomes the focus for conflict with the teacher. Instructional time is consumed in confrontation and behavior management rather than content and skill-related interactions. (Teacher expectations and student performance/group norms)

## REFERENCES

Adams, G. R. Racial membership and physical attractiveness effects on preschool teacher's expectations. *Child Study Journal*, 1978, *8*, 29-41.

Ames, C., Ames, R., & Felker, D. W. Effects of competitive reward structure and valence of outcome on children's achievement attributions. *Journal of Educational Psychology*, 1977, *69*, 1-8.

Andrews, G. R., & Debus, G. R. Persistence and causal perception of failure: Modifying cognitive attributions. *Journal of Educational Psychology*, 1978, *70*, 154-66.

Arlin, M., & Whitley, T. W. Perceptions of self-managed learning opportunities and academic locus of control: A causal interpretation. *Journal of Educational Psychology*, 1978, *70*, 988-92.

Bandura, A. *Social learning theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1977.

Bergan, J. R. *Behavioral consultation*. Columbus, Ohio: C. E. Merrill, 1977.

Bergan, J. R. & Henderson, R. W. *Child development*. Columbus, Ohio: C. E. Merrill, 1979.

Bergan, J. R., & Parra, E. Variations in IQ testing and the letter-learning and achievement of Anglo and bilingual Mexican American children. Unpublished manuscript, University of Arizona, 1978.

Blackwell, J. E. *The black community: Diversity and unity*. New York: Dodd, Mead, 1975.

Bloom, B. S. *Human characteristics and school learning*. New York: McGraw-Hill, 1976.

Brophy, J. S., & Good, T. *Teacher-student relationships: Causes and consequences*. New York: Holt, Rinehart & Winston, 1974.

Byalick, R., & Bersoff, D. N. Reinforcement practices of black and white teachers in integrated classrooms. *Journal of Educational Psychology*, 1974, *66*, 473-80.

Byers, P., & Byers, H. Non-verbal communication in the education of children. In C. Cazden, V. Joh, & D. Hymes (Eds.), *Functions of language in the classroom*. New York: Teachers College Press, 1972.

Covington, M. V. & Omelich, C. L. Effort: The double-edged sword in school achievement. *Journal of Educational Psychology*, 1979, *71*, 169-82.

Diener, C. I., & Dweck, C. S. An analysis of learned helplessness: Continuous changes in performance, strategy, and achievement cognitions following failure. *Journal of Personality and Social Psychology*, 1978, *36*, 451-62.

Doebler, L. K., & Eicke, F. J. Effects of teacher awareness of the educational implications of field-dependent/field independent cognitive style on selected classroom variables. *Journal of Educational Psychology*, 1979, *71*, 226-32.

Dollinger, S. J., & Laub, S. I. The interaction of locus of control expectancies and providing purpose on children's motivation. *Journal of Research in Personality*, 1977, *11*, 118-27.

Dusek, J. B. Do teachers bias childrens' learning? *Review of Educational Research*, 1975, *45*, 661-84.

Dusek, J. B., & O'Connell, E. J. Teacher expectancy effects on the achievement test performance of elementary school children. *Journal of Educational Psychology*, 1973, *65*, 371-77.

Dweck, C. S. The role of expectations and attributions in the alleviation of learned helplessness. *Journal of Personality and Social Psychology*, 1975, *25*, 109-16.

Elashoff, J., & Snow, R. E. (Eds.). *Pygmalion revisited*. Worthington, Ohio: C. A. Jones, 1971.

English, R. Beyond pathology: Research and theoretical perspectives on black families. In L. E. Gary (Ed.), *Social research and the black community: Selected issues and priorities*. Washington, D.C.: Institute for Urban Affairs and Research, Howard University, 1974.

Farris, B. E., & Glenn, N. D. Fatalism and familism among Anglos and Mexican Americans in San Antonio. *Sociology and Social Research*, 1976, *60*, 393-402.

Foster, G., & Ysseldyke, J. Expectancy and halo effects as a result of artificially induced teacher bias. *Contemporary Educational Psychology*, 1976, *1*,37-45.

Frieze, I., & Weiner, B. Cue utilization and attributional judgments for success and failure. *Journal of Personality*, 1971, *39*, 591-606.

Gallimore, R. Affiliation motivation and Hawaiian-American achievement. *Journal of Cross-Cultural Psychology*, 1974, *5*, 481-91.

Gillung, T. B., & Rucker, C. N. Labels and teacher expectations. *Exceptional Children*, 1977, 43, 464-65.

Good, T. L., & Brophy, J. E. Changing teacher and student behavior: An empirical investigation. *Journal of Educational Psychology*, 1974, 66, 390-405.

Grebler, L., Moore, J. W., & Guzman, R. C. *The Mexican American people: The nation's second largest minority*. New York: Free Press, 1970.

Griffith, B. *American me*. Boston: Houghton-Mifflin, 1948.

Hawkes, G. R., & Taylor, M. Power structure in Mexican and Mexican American farm labor families. *Journal of Marriage and the Family*, 1975, 37, 807-11.

Heller, C. S. *Mexican American youth: Forgotten youth at the crossroads*. New York: Random House, 1966.

Henderson, R. W. Personal and social causation in the school context. In J. Worrell (Ed.), *Developmental psychology for education*. New York: Academic Press, 1980.

Henderson, R. W. & Hennig, H. Relationships among cooperation-competition and locus of control in academic situations among children in traditional and open classrooms. *Contemporary Educational Psychology*, 1979, 4, 121-31.

Henderson, R. W.,& Merritt, C. B. Environmental backgrounds of Mexican American children with different potentials for school success. *Journal of Social Psychology*, 1968, 75, 101-106.

Holtzman, W. H., Diaz-Guerrero, R., & Swartz, J. D., in collaboration with L. Lara-Tapia, L. M. Laosa, M. L. Morales, I Reyes Lagunes, & D. B. Witzke. *Personality development in two cultures: A crosscultural longitudinal study of school children in Mexico and the United States*. Austin: University of Texas Press, 1975.

Kagan, S., & Buriel, R. Field-dependence-independence and Mexican American culture and education. In J. L. Martinez (Ed.), *Chicano psychology*. New York: Academic Press, 1977.

Kifer, E. Relationships between academic achievement and personality characteristics: A quasi-longitudinal study. *American Educational Research Journal*, 1975, 12, 191-210.

Kirk, S. A., & Gallagher, J. J. *Educating exceptional children* (3rd ed.). Boston: Houghton-Mifflin, 1979.

Klein, S. S. Student influences on teacher behavior. *American Educational Research Journal*, 1971, 8, 403-22.

Kluckhohn, F. R., & Strodtbeck, F. *Variations in value orientations*. New York: Row, Peterson, 1961.

Kroeber, A. L., & Kluckhohn, C. *Culture: A critical review of concepts and definitions*. New York: Vintage Books, 1952.

Laosa, L. M. Inequality in the classroom: Observational research on teacher-student interactions. *Aztlan International Journal of Chicano Studies Research*, 1978, 8.

Laosa, L. M. Social competence in childhood: Toward a developmental, socio-culturally relativistic paradigm. In M. W. Kent & J. E. Rolf (Eds.), *Primary prevention of psychopathology*. Vol. III: *Social competence in children*. Hanover, N.H.: University Press of New England, 1979.

Laosa, L. M., & DeAvila, E. A. Development of cognitive styles among Chicanos in traditional and dualistic communities. *International Journal of Psychology*,in press.

Lefcourt, H. M. *Locus of control: Current trends in theory and research*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1976.

Lewis, D. Education and sense modality: Class and culture conflict in the schools. Unpublished manuscript, Oakes College, University of California, Santa Cruz, no date.

Lewis, O. *Children of Sanchez*. New York: Random, 1961.

Lewis, O. *Five families*. New York: Basic Books, 1959.

Lewis, O. *La Vida: A Puerto Rican family in the culture of poverty*. New York: Random, 1966.

Lilly, M. S. (Ed.). *Children with exceptional needs*. New York: Holt, Rinehart & Winston, 1979.

Lockheed, M. E. Some causes and consequences of teacher expectations concerning pupil performance. *California Journal of Teacher Education*, 1977, 4, 40-56.

Luce, S. R., & Hoge, R. D. Relations among teacher rankings, pupil-teacher interactions, and academic achievement: A test of the teacher expectancy hypothesis. *American Educational Research Journal*, 1978, 15, 489-500.

Madsen, W. *The Mexican American of South Texas*. New York: Holt, Rinehart, & Winston, 1964.

Mercer, J. Sociological perspectives on mental retardation. In H. C. Haywood (Ed.), *Social-cultural aspects of mental retardation: Proceedings of the Peabody NIMH Conference*. New York: Appleton-Century-Crofts, 1970, 378-91.

Mercer, J. *Labeling the mentally retarded: Clinical and social system perspectives on mental retardation*. Berkeley: University of California Press, 1973.

Montiel, M. The Chicano family: A review of research. *Social Work*, March 1973, 22-31.

Moynihan, D. P. *The Negro family: The case for national action*. Washington, D.C.: U.S. Department of Labor, March, 1967.

Parra, E., & Henderson, R. W. Mexican American perceptions of parent and teacher roles in child development. *Bilingual Review*, 1977, 4, 210-17.

Parsons, T. The social structure of the family. In R. N. Anshen (Ed.), *The family: Its function and destiny*. New York: Harper & Row, 1949.

Paz, O. *The labyrinth of solitude: Life and thought in Mexico*. New York: Grove Press, 1961.

Pettigrew, T. *A profile of the Negro American*. Princeton, N.J.: D. VonNostrand, 1964.

Ramirez, M. Cognitive styles and cultural democracy in education. *Social Science Quarterly*, 1973, 53, 895-904.

Ramirez, M., III & Castaneda, A. *Cultural democracy, bicognitive development, and education*. New York: Academic Press, 1974.

Rhodes, W. C. The disturbing child: A problem of ecological management. *Exceptional Children*, 1967, 33, 449-55.

Rosenthal, R., & Jacobsen, L. *Pygmalion in the classroom*. New York: Holt, Rinehart & Winston, 1968.

Ross, M. B., & Salvia, J. Attractiveness as a biasing factor in teacher judgments. *American Journal of Mental Deficiency*, 1975, 80, 96-98.

Sanders, M., & Scholz, J. P. Field independence match-mismatch and school achievement gains of Anglo American and Mexican American children. Paper presented at the Western Psychological Association Meetings, Los Angeles, April 1976.

Thomas, A. Learned helplessness and expectancy factors: Implications for research in learning disabilities. *Review of Educational Research*, 1979, *49*, 220-21.

Weisz, J. R. Perceived control and learned helplessness among mentally retarded children: A developmental analysis. *Developmental Psychology*, 1979, *15*, 311-19.

Williams, T. Teacher prophecies and the inheritance of inequality. *Sociology of Education*, 1976, *49*, 223-36.

Willis, S., & Brophy, J. Origins of teachers' attitudes toward young children. *Journal of Educational Psychology*, 1974, *66*, 520-29.

Winetsky, C. S. Comparisons of the expectations of parents and teachers for the behavior of preschool children. *Child Development*, 1978, *49*, 1146-54.

Witkin, H. A. A cognitive style approach to cross-cultural research. *International Journal of Psychology*, 1967, *2*, 233-50.

Witkin, H. A., & Goodenough, D.R. Field dependence and interpersonal behavior. *Psychological Bulletin*, 1977, *84*, 661-89.

Witkin, H. A., Goodenough, D. R., & Oltman, T. K. Psychological differentiation: Current status. (Educational Testing Service Research Bulletin 77-17). Princeton, N.J.: Educational Testing Service, 1977.

Woods, F. J. *Cultural values of American ethnic groups*. New York: Harper, 1956.

## RECOMMENDED READINGS

Brophy, J. E., & Good, T. *Teacher-student relationships: Causes and consequences*. New York: Holt, Rinehart & Winston, 1974.

This book presents an extensive review of research on teacher expectancies and variations in teacher-pupil interaction patterns. Relations between interaction patterns and student characteristics are presented. Successful approaches for changing interaction patterns that bias the learning opportunities of children are described.

Cordasco, R., & Bucchioni, E. *The Puerto Rican experiences: A sociological sourcebook*. Totowa, N.J.: Rowman & Littlefield, 1973.

This sourcebook may be used as an alternative to J. M. Gallardo. It provides information on the island background and the migration of Puerto Ricans to the mainland. A separate section deals with repression and resistance, and with conflict and acculturation processes on the mainland. The section, "Education on the Mainland," is especially relevant to the concepts presented in this section of the module.

English, R. A. Socialization and black family life. In L. E. Gary (Ed.), *Social research and the black community: Selected issues and priorities*. Washington, D.C.: Institute for Urban Affairs and Research, Howard University, 1974, 39-52.

Gallardo, J. M. (Ed.). *The Puerto Rican experience (Proceedings of the conference on education of Puerto Rican Children on the Mainland, Oct. 18-21, 1970). New York: Arno Press, 1975.*

*This slim volume contains a collection of papers on educational problems faced by Puerto Ricans on the mainland. The reactions of conference participants to each presentation are included. Language problems and the need for bilingual education is a major theme. Presentations on the cultural background of Puerto Rican children by Juan J. Maunez and Carmen Silva Garcia are of particular relevance to this module segment.*

Lightfoot, S. L. *Worlds apart: Relationships between families and schools*. New York: Basic Books, 1978.

This book may be used as an alternative to the chapter by English. In this book a chapter, "Boundaries and Bridges," describes the conflict between schools and black families. Concepts from structural sociology (e.g., role definitions and expectations) are called upon to describe home-school discontinuities. Due attention is also given to discrepancies between theoretical conceptions of Parsonian sociologists and the real environments of children. Chapter 4, "Black Dreams and Closed Doors," challenges the assumption that blacks have low educational aspirations with historical information on the importance blacks have attributed to education. See the review of this book by R. W. Henderson, *Journal of School Psychology*, 1980 18(1).

Montiel M. The Chicano family: A review of research. *Social Work*, March 1973, 22-31.

This article presents a critical examination of research on the Chicano family, the "*machismo* cult," socialization processes, and interventions designed to remedy presumed deficiencies in the functioning of Chicano individuals and their families.

# NONBIASED ASSESSMENT: UNDERSTANDING LANGUAGE CHARACTERISTICS

Joseph H. Matluck
*The University of Texas at Austin*

Betty J. Mace-Matluck
*Southwest Educational Development Laboratory*
*Austin, Texas*

# Table of Contents

- Introduction ........ 57
- Objectives ........ 58
- Pretest ........ 59
- The Nature of Language ........ 61
  - Subsystems of Language ........ 61
  - Language as a Reflection of Culture ........ 62
  - Language as a Vehicle of Communication ........ 62
  - Language Variation ........ 63
  - Language Acquisition ........ 63
- Assessing Language ........ 65
  - Purpose of Language Testing ........ 65
  - Determining Language Status ........ 66
- Characteristics of Oral Language Proficiency Tests ........ 67
  - Types of Tests ........ 67
  - Modes of Testing ........ 67
  - Content ........ 67
  - Psychometric Qualities ........ 68
- Legally-Mandated Assessment and Intervention ........ 68
  - Identification of Students of Limited English-Speaking Ability (LESA) ........ 68
  - Identification of Limited English Proficient (LEP) Students ........ 69
- Diagnosing Learning Difficulties of Limited English-Speaking and Bilingual Children ........ 69
- Posttest ........ 69
- Simulations ........ 71
- Pretest Key ........ 75
- References ........ 75
- References ........ 75
- Annotated Bibliography ........ 76

## Introduction

In the past 10 years, language assessment has become of increasing concern to school district educators. This concern first arose out of a change in our immigration laws that allowed a steady influx of new immigrants into the country from non-English-speaking areas, beginning in the 1960s and increasing in tempo during the 1970s. This influx also produced an awakening in some populations—largely Mexican Americans and Puerto Ricans—that their children were almost as severely handicapped in school by limited proficiency in English as were the immigrant children. The need grew for determining the language status of these limited-English-speaking children. Interest in assessment instruments and in the linguistic rationale on which they are based also grew. The Civil Rights Act of 1964, the Bilingual Education Act in 1968, and the 1974 *Lau* decision in *Lau vs. Nichols* (1974) heightened and intensified this interest.

State and federal guidelines which were prepared to assist school districts in complying with the court decisions and public laws, require school districts to assess the English language skills of children from non-English-speaking backgrounds, to classify them on the basis of language status, and to provide instructional programs commensurate with their language and academic needs. In keeping with the immediate needs of school districts, as defined in terms of the *Lau* decision and P.L. 94-142, the primary focus of this section is on assisting school personnel to classify children for the purpose of placement in appropriate educational programs. The importance of obtaining information for the purpose of diagnosing and developing interventions for individual students is recognized. Although interventions are not the major focus in this section, information presented here may be useful as a foundation for acquiring further insights into the language and academic needs of bilingual children.

Because assessment in schools generally falls within the purview of the school psychologist, other school personnel increasingly have looked to them for guidance and assistance in conducting language assessments. Although the role of the school psychologist may not include direct administration of language tests, knowledge about language and general assessment principles enables the psychologists to work with language specialists and, when necessary, to provide direction to other school personnel. Thus, in view of the recent emphasis on language testing, school psychologists must take an active role in language assessment and should acquire as much knowledge and insight into this process as possible.

Included in this section of the module, therefore, are (a) a basic discussion of the nature of language and of the language acquisition process; (b) an overview of commonly used procedures for assessing language skills of children from non-English-speaking backgrounds; (c) general characteristics of present-day oral language proficiency tests; and (d) a summary of legally required language assessment and intervention procedures.

70

## OBJECTIVES

Upon completion of this section of the module, the participant should be able to perform the following:

1. Assist school personnel in understanding the language behavior of limited English-speaking and bilingual children.
2. Assist school districts in the evaluation and selection of language assessment instruments.
3. Use data from language assessment to assist in the diagnosis of learning difficulties and in the development of intervention strategies for limited English proficient (LEP) and bilingual students.
4. Discuss and define in general terms the current requirements of legally required assessment and intervention procedures for children with non-English-speaking home backgrounds.

## PRETEST

1. For the school psychologist two of the most important areas in the field of language are ______
   (a) the history of recent immigration; civil rights legislation
   (b) procedures for assessing language; the nature of language and the acquisition process
   (c) an understanding of minority parents; a thorough knowledge of their language
   (d) local school policy regarding testing; knowledge of the inventory of tests used locally
2. Language is, as defined by linguists, ______
   (a) a system of articulated sounds used by any species
   (b) a system of automatic vocal and written symbols used for communication
   (c) a unique reflection of a particular culture
   (d) the only way groups can communicate
3. The following is a valid statement about language: ______
   (a) it is not uniquely human
   (b) it does not perform a social function
   (c) it is not the same as writing
   (d) it is not systematic
4. In the sentence, *She take to her seven buck for him read*, as spoken by a nonnative English speaker, the speaker has made errors of ______
   (a) phonology, morphology, and syntax
   (b) phonology and syntax, but not of morphology
   (c) phonology and morphology, but not of syntax
   (d) morphology and syntax, but not of phonology
5. In the preceding sentence ______
   (a) the full semantic import of the speaker's message would be transmitted
   (b) there will be no ambiguities in the transmission of the semantic import of the speaker's message
   (c) only a portion of the semantic import of the speaker's message would be transmitted
   (d) the listener will have no doubts about the semantic import of the message being transmitted
6. One language is not "better" than another; it is just that ______
   (a) some languages are more complicated than others
   (b) some languages are more descriptive than others.
   (c) the language one doesn't know just doesn't have the words to express what one's native language does so simply
   (d) the language one knows (one's native language) seems better able to handle the necessities of communication
7. Because language is one of the major components of culture, it reflects ______
   (a) the way that a society views the reality around it.
   (b) the quality and quantity of literature written by a society
   (c) the superior or inferior way of thinking of a society
   (d) the more effective or less effective ways that a society has of expressing itself
8. The primary linguistic system is ______
   (a) the written language: reading and writing
   (b) the kinesic: touching and feeling
   (c) the spoken language: understanding and speaking
   (d) the paralinguistic: grunts and groans
9. The most important difference between the spoken and written language is that ______
   (a) writing is not as important as speaking either to the individual or to the species
   (b) the spoken language is an imperfect and less precise substitute for the written language
   (c) the written language is quite different from the spoken language
   (d) the spoken language is not as symbolic as the written language.
10. The notion that the written language is a code that can be decoded with equal ease regardless of oral proficiency is ______
    (a) a mistaken one, based on the belief that the real language is the written language and the spoken language

merely an imperfect substitute
(b) a mistaken one, based on the belief that the real language is the spoken language and the written language merely an imperfect substitute
(c) a valid one, based on the belief that the written language must conform to the rules of the spoken language
(d) a valid one, based on the belief that the spoken language must conform to the rules of the written language

11. Of the four language skills, the receptive (or passive) skills are ______
(a) reading and speaking
(b) understanding and writing
(c) speaking and writing
(d) understanding and reading

12. Of the four language skills, the expressive (or active) skills are ______
(a) reading and speaking
(b) understanding and writing
(c) speaking and writing
(d) understanding and reading

13. Of the primary and secondary systems on the one hand, and the receptive and expressive skills on the other, ______ are normally learned earlier by native speakers.
(a) the primary system and receptive skills
(b) the primary system and expressive skills
(c) the secondary system and receptive skills
(d) the secondary system and expressive skills

14. The notion that a dialect is an "impure," imperfect attempt to emulate the "standard" language is ______
(a) a mistaken one, because most people speak at least two dialects
(b) a mistaken one, because people with less education don't speak the standard language
(c) a valid one, because people with more education do speak the standard language.
(d) a valid one, because people with less education cannot communicate as well as those with more education

15. An idiolect is ______
(a) usually associated with standard dialects
(b) usually associated with nonstandard dialects
(c) both social and regional
(d) neither social or regional

16. A speaker's style or register is determined ______
(a) by the variable of education
(b) by the variable of region
(c) by the variable of whim or notion
(d) by many variables

17. The dialect arbitrarily designated by a society as "standard" ______
(a) is usually the best and most expressive dialect
(b) is usually socio-economically determined
(c) is usually based on the choice of the most prestigious dialect
(d) is usually stylistically determined

18. In the acquisition by monolingual children of their native language ______
(a) we cannot predict their stages of language development on the basis of age
(b) simple imitation of adult speech assures immediate acquisition of most structures
(c) most of the basic structures of the spoken language are acquired by the age of 5-6
(d) there is little lexico-semantic development after age 6

19. Second-language acquisition differs from first-language acquisition in that ______
(a) developmental processes related to developing brain chemistry play an important part in second-language acquisition
(b) the first language is usually fairly well established at the time of acquisition of the second
(c) motivation and opportunity are more important in first-language acquisition than in second
(d) the optimal age for acquiring a second language is well established

20. Results of general school achievement tests may not be reliable and valid for a particular bilingual or limited English proficient student ______
(a) if the test content is linguistically and culturally appropriate
(b) if the test is in the student's own dialect or language
(c) if the test administrator is familiar to the student
(d) if the test is in the student's weaker language

21. Language testing is of concern to school psychologists because ______
(a) they are often bilingual
(b) they are usually experts in language
(c) language and learning difficulties are often related
(d) language and instructional planning are unrelated

22. The population to be tested and the type of language assessment instrument to be selected are largely dependent upon ______

(a) the projected use to be made of the results
(b) availability of bilingual staff
(c) local option
(d) the cost of the available instruments

23. Measures designed to determine students' strengths and weaknesses in a particular language are ______
(a) aptitude tests
(b) placement tests
(c) achievement tests
(d) diagnostic tests

24. When we look at the linguistic skills of bilingual children in most U.S. elementary schools, we find that ______
(a) most of them are fluent speakers of their home language
(b) they are usually not a homogeneous group linguistically
(c) most of them have native-like control of both languages
(d) they have usually retained little of their home language ability

25. The legislation mandating school districts to define the language status of each child has resulted in wide-scale language testing that has as its prime purpose ______
(a) measurement of students' achievement
(b) diagnosis of students' language strengths and weaknesses
(c) determination of students' language aptitude
(d) placement of students into appropriate programs

26. Recent research has shown that one of the positive effects associated with bilingualism may be ______
(a) accelerated cognitive growth
(b) cognitive inflexibility
(c) higher school achievement
(d) greater intellectual achievement

27. An example of a direct measure for determining the language dominance of bilingual children is ______ in each language.
(a) fluency tests
(b) self-rating scales
(c) oral proficiency tests
(d) word association tests

28. Instruments like standardized tests of oral proficiency can measure directly and independently a child's oral performance in each of her/his languages. A distinct advantage of such measures is that ______.
(a) they are usually time consuming to administer and score
(b) they must sample a variety of language functions
(c) they must be comparable
(d) they can measure competence in each of the four skills

29. Discrete-item tests ______
(a) are usually conversation based
(b) usually contain a specific number of items
(c) usually attempt to elicit "free speech"
(d) often take the form of a structured interview

30. Answer ______ is a true statement about test modes which are used to assess production skills.
(a) The Structured Response mode requires the child to supply a complete answer to a direct question
(b) In the Sentence Repetition mode the child must be able to decode as well as encode the material to be repeated
(c) Listening comprehension is not required in the "Retell" part of the Oral Composition (Tell; Retell) mode
(d) The Direct Questioning mode uses very tightly controlled cues

31. With respect to the content of oral proficiency tests, ______
(a) all tests attempt to assess all four subsystems of language
(b) all tests attempt to assess phonology
(c) some tests attempt to score only syntax
(d) few tests attempt to assess morphology

32. In view of the severe limitations on currently available oral proficiency tests with respect to their psychometric qualities, ______
(a) there is little justification for continuing to use them
(b) they should be replaced by subjective measures alone
(c) currently available statistical procedures should no longer be used for reporting language data
(d) test scores should be used with extreme caution

33. Of the currently available oral proficiency tests, ______
(a) most are standardized on an adequate number of children
(b) very few have demonstrated that they meet acceptable psychometric standards
(c) most have gone through a complete development and refinement process
(d) very few attempt to use statistical procedures

34. The necessity for school districts to determine each child's language status objectively was directly mandated by ______
    - (a) the Lau Decision
    - (b) the Civil Rights Act
    - (c) the U.S. Supreme Court
    - (d) the Lau Remedies
35. The major problem with the LESA classification scheme was that ______
    - (a) it was based on oral proficiency alone
    - (b) it was based on language dominance alone
    - (c) it mandated special programs for all children
    - (d) it mandated no special programs at all
36. The LEP scheme classifies children on the basis of ______
    - (a) their degree of home-language proficiency
    - (b) their degree of English-language proficiency
    - (c) their degree of English dominance
    - (d) subjective judgments by teachers
37. Inasmuch as so many measures for diagnosing learning disabilities depend upon a child's language abilities, it would seem advisable ______
    - (a) to translate all measures into the child's home language
    - (b) to eliminate language difficulties as the possible source of the child's learning problems
    - (c) to administer a language dominance test to all bilingual children at the beginning of the school year
    - (d) to administer a language aptitude test to all children
38. Diagnostic language tests for bilingual children can be of value to the school psychologist ______
    - (a) if they are comparable in their structure and content
    - (b) if they test only English
    - (c) if they test only the home language
    - (d) if they are administered in her/his presence

## The Nature of Language

*As defined by most linguists, the term "language" refers to spoken language only.* True______ False______

Language has been defined in many ways. For some scholars it is a system of articulated sounds organized by human thought and used by a group of humans for the purpose of communication. For others, it is a system of arbitrary vocal symbols through which human beings belonging to a particular social group communicate, interact, and cooperate. These symbols operate automatically as a set of habits which are fully developed and controlled in both production and reception by the native-born adult members of that group. To some people, language is a gift of God only to the human species; to others, it is the unique reflection of a particular culture and the ways in which that group views, interprets, and expresses the reality that it experiences daily through the senses.

Common to all views of language is that it (a) is uniquely human; (b) is systematic; (c) is vocal; (d) performs a social function; and (e) is not the same as writing or other forms of communication (English Language Services, 1968; Fromkin & Rodman, 1974; Saville-Troike, 1976).

### Subsystems of Language

*Define briefly but concisely the four subsystems of language. For discussions of them see Fromkin & Rodman, 1974, pp. 9-12; Kehoe, 1971, pp. 5-17; Langacker, 1968, pp. 21-42.*

Linguists analyze language in different ways; thus, no one analysis—perhaps not even a single definition—can satisfy everyone. We offer here one view of the subdivision of language into four subsystems: phonology, morphology, syntax, and semantics.

The *phonology* of a language is the subsystem that deals with the sounds of the language, their production, the perception of them, their combinatory possibilities, and the ways in which they function in distinguishing meaning. That is, *pit* is distinguished by English native speakers from *bit* by the fact that /p/ and /b/ are members of different sound families, or phonemes, and *through* is distinguished from *throw* for the same reason relating to the phonemes /u/ and /o/.

The *morphology* of a language is the subsystem that deals with the forms which signal meaning in the language, whether they be words or parts of words. These forms are called morphemes. Most words are lexical morphemes, that is, they usually signal lexical or vocabulary meaning (e.g., *book, tie*). Parts of words most often are grammatical morphemes, that is, they usually signal grammatical meaning. For example, the /s/ of *books* signals more than one book, and /s/ of the verb *ties* and the /d/ of *tied* give the listener information on the grammatical tense or person.

The *syntax* of a language is the subsystem by which the morphemes are arranged into the larger units of phrases, clauses, and sentences. The syntactical rules of English, for example, tell us that *Johnny bit Mary* is a permissible sentence, and that it means that not only did something take place in the past (morphology: past tense of the verb *bite*) but, also, that Johnny was the doer and Mary the receiver of the biting. *Mary bit Johnny* reverses the order of the subject and object. This does not mean that the same syntactical rules can be applied to other languages. In fact, many languages—Spanish among them—can put the object first, providing that certain other rules are followed (e.g., A Maria le mordio Juanito = Johnny bit Mary). In English we cannot say *Mary John bit, Bit John Mary, John bite Mary,* or *John bitten Mary*, but there are languages in which one or more of these forms might be perfectly permissible. In yet another area of syntax, English syntactical rules allow us to say *I know the professor*, but not usually *I know the Professor Jones*. Not so again with Spanish, for instance: *Conozco al profesor Garcia* . It readily can be seen from the examples that both the morphological and syntactical rules of a language are intimately related and intertwined. In fact, many linguists refer to the morphosyntactic rules of a language as the rules governing its grammatical structure.

The *semantic* structure of a language is the subsystem that (a) relates sound to meaning or vice versa, or, from another point of view, (b) deals with the conceptual import of the message which speakers want to communicate through language and the ways in which they manipulate the rules of each of the other three subsystems so that the listener receives the exact message which the speakers had conceptualized in the first place. *He took to her right away* is not the same as *She took him right away*. Obviously, if an error is made in phonology, morphology, syntax, or the selection of the appropriate lexical items, than one cannot be certain that the listener will receive exactly the same message which the speaker intended. *He took to her right away* is vastly different in meaning from *He talked of her right of way*. The phonological differences in these two utterances are very slight, but they produce great lexical and grammatical differences. When both speaker and listener are native speakers, even the most minimal differences in the speaker's utterances are detected and decoded by the listener with exactly the same conceptual nuances which the speaker intended. If, however, one of the two does not have native-like control of one or more of the rules of even one subsystem, some kind of breakdown or lag in communication will occur because of the listener's greater or lesser loss of meaning.

Another example of how semantic systems work involves idioms (i.e., phrases or utterances whose meaning is not the simple, literal total of their component parts). For example, *She kicked the bucket* refers neither to a pail nor the act of kicking; and *He's got ants in his pants* refers neither to insects nor trousers. Because native speakers know the semantic rules, they know which utterances to interpret literally and which, idiomatically.

Syntactic and semantic sytems are very complex indeed. The information just presented is intended to give only a working notion of both systems within the framework of the four subsystems.

### Language as a Reflection of Culture

---

*English is much better at describing an occurrence than most other languages.* *True______ False______*

---

Cultures, as defined by most linguists and anthropologists, are "structured systems of patterned behavior" (Lado, 1957, pp. 110-123) or, to put it another way, language is "the expressive dimension of culture" (Saville-Troike, 1976, pp. 27-29, 45-67). Children who grow up in a particular society or culture learn a whole complex set of behavior patterns which are peculiar to that culture, and they learn to look at and interpret the daily reality around them in terms of those learned patterns. At the same time, they learn the language system which is used in that culture to label, arrange, and interpret the reality in a way that is peculiar to that language and that culture alone. This is because language is one of the major components of culture; it is true in every area of language. On the phonological level, the language has sounds, combinations of sounds, contrasts of sounds, and so forth, which are peculiar to that language and culture. On the morphological level, grammatical signals fall into certain categories and types peculiar to that language and culture. And the same is true on the lexical, syntactic, and semantic levels. For example, (a) the English word *snow* (both noun and verb) is divided into as many as seven different categories in certain Eskimo languages; (b) the color spectrum in some languages is divided up differently than is ours; (c) Czech has sound combinations such as /ctvrt/ *(tchtvrt)*, with no intervening vowel sounds, which are impossible for English or Spanish; (d) an English speaker says, *I dropped my pencil*, with no necessity, except under extraordinary conditions, to specify whether it was dropped accidentally or purposefully; a Spanish speaker, on the other hand, says, *Se me cayo el lapiz*—the pencil, associated with me (as mine) dropped itself (was dropped, fell down somehow). In both languages the usual meaning of the sentence is that it was accidental; yet the Spanish construction shows the accidental nature of the event whereas the English structure is ambiguous. (e) In Spanish one says, *Se puso el sol*, and *Se puso bravo*, using the same verb put in the reflexive *(se puso)* for the sun's *setting* and somebody's *becoming* or *getting* angry. Spanish speakers look at the reality of the sun setting as "putting itself" over the horizon, and the reality of becoming angry as "putting oneself" in the state of anger. There are literally thousands of such syntactic examples in every language. To dismiss them simply as "idioms" is to miss the main point: They are different grammatical structures because each culture views the reality (the dropping of the pencil, the setting of the sun, and the becoming angry) in different ways, with difference interpretations of the reality, and, therefore, with different ways (grammatical structure) of expressing or describing that reality.

### Language as a Vehicle of Communication

---

*Non-native speakers of English can learn to read English as easily as native English speakers regardless of their proficiency in spoken English.* *True______ False______*

---

**Primary and secondary systems.** The primary vehicle by which people communicate is spoken language. Accompanying this primary vehicle are the nonverbal aspects of communication: the kinesic (e.g., gestures), which is neither vocal nor verbal, and the paralinguistic (grunts, groans, etc.), which is vocal but not verbal. Nonverbal communication is an important but a secondary vehicle. Figure 1 illustrates how the primary system functions.

Writing is not language. It is a system of communication and is *based* on language, but it is not language in and of itself. It is a substitute for language. It does not mirror speech exactly and the number of written symbols is most inadequate for indicating not only the huge number of different sounds in any language but, also, the suprasegmental elements, such as stress, pitch, juncture, and rhythmic and intonational features.

Figure 1. The speech chain: Primary linguistic system. Source: W. G. Moulton, Linguistics, *NEA Journal*, January, 1965.

Written language is important, however. Children must become literate if they are to be able to read books, and they must learn to write as a basic prerequisite for being able to submit compositions, reports, take examinations, and so on, all of which also are necessary to the educational process and beyond.

The written language is important but *different* from the spoken language and usually is a substitute for it. We use written language only when communication through the primary vehicle is impossible, undesirable, unfeasible, or inconvenient.

The preceding ideas are important, especially for teachers.

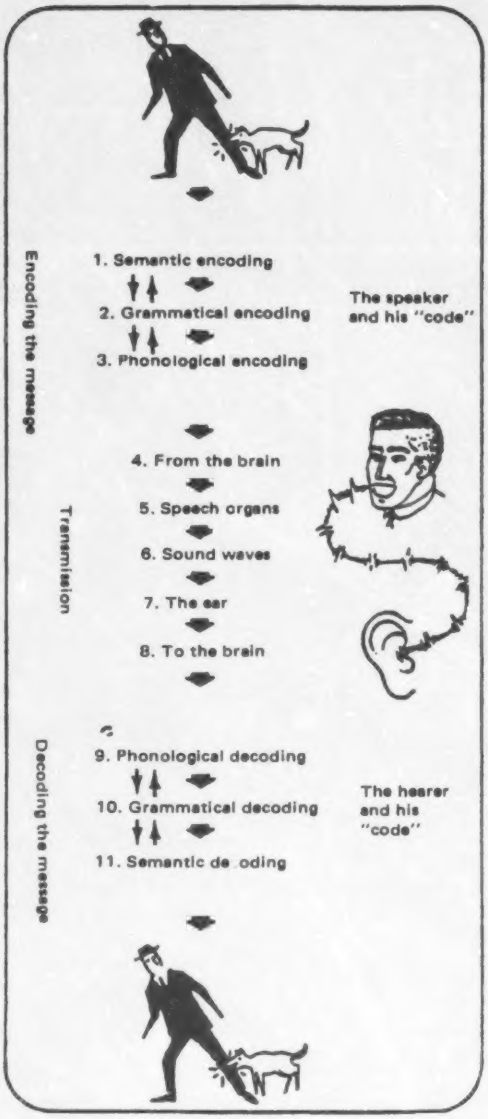


Figure 1. The speech chain: Primary linguistic system.
Source: W.G. Moulton, Linguistics, NEA Journal, January, 1965.

Our constant preoccupation with and concentration on teaching children to read tends to distort our criteria and to reverse reality to the point where we believe that the written is the real language and the spoken language is only an imperfect substitute. In turn, this misconception leads us to the mistaken judgment that the norms of the written language also govern the spoken language and that rules for the latter must conform to those of the former.

The same reversal of reality in the area of reading tends to give us the mistaken notion that the written language is a code which can be decoded with equal ease by native and non-native children alike (see English Language Services, 1968; Fromkin & Rodman, 1974; Moulton, 1970).

---

*List the* **active** *language skills.*

---

**Receptive and expressive skills.** If we examine the four most frequently discussed language skills—understanding, speaking, reading, and writing—we find that understanding and speaking make up the *primary* linguistic system that comprises the spoken language. Reading and writing make up the secondary system, which comprises the written language. From another point of view, understanding and reading are the *passive* skills and require receptive and perceptive abilities (decoding skills), whereas speaking and writing are the *active* skills and require production abilities (encoding skills).

The primary skills of understanding and speaking generally are learned by native speakers at a younger age than are the secondary skills of reading and writing. Also, the passive skills generally develop at a more rapid rate than the native speakers can produce, and they can read more than they can write.

**Language Variation**

---

*Define the terms* **idiolect, dialect,** *and* **standard language.**

---

Myths abound in this area (e.g., "I speak the 'pure,' 'real,' standard language, you-all speak a dialect—an 'impure,' imperfect attempt to emulate my standard"). In fact, of course, everybody speaks a dialect or, rather, everybody speaks her/his own idiolect (i.e., her/his variety of language with idiosyncracies of speech that make one idiolect different from all others). But each idiolect is very similar to the idiolects of the other people who live in the same region. All speak the same regional dialect of the language. In certain regions, for example, *ten* and *tin* are said with the same vowel sound and are phonologically not differentiated, but *cot* and *caught* are pronounced with different, distinguishable, and contrasting vowel sounds; in other regions the reverse is true; in yet another, *cot, caught, ten* and *tin* are all phonologically differentiated but *cot* and *cart* are not (the so-called *r*-less dialects), as is also the case with *hominy* and *harmony*. These four words may be perfectly contrasted in the other dialects we mentioned. Some say *pail*, others *bucket;* some say *frying pan*, others *skillet*.

Most people speak more than one dialect. They have a regional dialect and a social dialect (which is largely determined by educational level). In addition, all speakers of a language have more than one *style* or *register*, depending on the immediate context or domain in which they are communicating. Simplistically, the latter can be divided into formal and informal, but it is more complicated than that. Some of the variables are the relationship of the speaker to her/his listener(s), the social context or domain (job, neighborhood, home, etc.), and the topic.

Finally, no dialect is, linguistically speaking, necessarily or inherently better, superior, more expressive, more logical, richer, more complex, more picturesque, and so on than any other dialect. Thus, judgments about the superiority or intrinsic value of one dialect over another arise from social, economic, or political and not linguistic factors. Given that the power structure of most societies usually lies with the higher socio-economic group, their dialect arbitrarily is designated as "standard" and prestigious, and all others are designated as "nonstandard" and less prestigious (see Saville-Troike, 1976; Framkin & Rodman, 1974; Dale, 1972).

**Language Acquisition**

---

*How does second-language acquisition differ from first-language acquisition? How is it similar?*

---

**First-Language Acquisition—Monolingualism.** Recent studies of first-language acquisition suggest that children the world over learn the native language through meaningful interactions with the people around them. Children are not "preprogrammed" to learn a particular language. They acquire the language that is most often spoken in their homes. Language acquisition is not the simple imitation of adult speech but a process in which children test hypotheses and gradually structure rules for the speech they hear. Children's language development proceeds through successive stages until the mismatch between what they hear and what they create is resolved. These stages are believed to coincide with certain maturational changes that are governed to a great extent by the physical development of the brain.

In monolingual development, children progress through stages that are predictable on the basis of age. By the time children enter school at age 5-6, they generally have entered the final stage of their native language acquisition. They have acquired most of the basic oral structures of the native language and have learned a good bit about the various styles of speaking in which people talk to each other under different circumstances, such as modifying speech in role playing and shifting styles when addressing authority figures.

The final stage of native language acquisition is of particular importance to the classroom teacher. Certain sounds may not be fully mastered by all children before age 8. A number of important syntactic structures (e.g., passive sentences and sentences containing the verbs *ask* and *promise*) are acquired between the ages of 5 and 10. Lexicosemantic development (the meaning of words) also occurs after age 6. Obviously, vocabulary is expanded and word meanings are elaborated throughout life. Although the preceding examples have been taken from English, there is convincing evidence that children the world over progress without the aid of formal instruction through successive stages in language development which are similar to those identified for English-speaking children.

**First-Language Acquisition—Bilingualism.** Many children living in various parts of the United States are exposed to and acquire two languages simultaneously in early childhood. Both are first languages for these children, although one is usually dominant in certain situations or with certain people. For example, if children hear one language from their parents and another from their playmates, they will tend to speak the home language with their parents and use the other with playmates. During early childhood, interaction with parents tends to be more extensive than that with playmates. Hence, the home language is likely to prevail as the child's dominant language. Perfect linguistic balance in communicating with parents and peers seems to be extremely difficult to achieve. Nonetheless, the relatively few studies of child bilinguals provide some evidence that children are capable of acquiring two or more languages simultaneously.

Case studies of children raised under bilingual conditions during the first three years of life, although offering contradictory evidence in certain instances, reveal several general points of agreement. Language acquisition follows the same developmental pattern in bilingual as in monolingual children. In the area of phonology (sound system of the language), some researchers have noted an initial period of confusion in learning the sounds of the two languages. When the exposure to the two languages is similar, the period of confusion is relatively short; when one language is dominant, the sound features of that language may be substituted for those of the weaker language. Similarly, words that are difficult to pronounce in one language may be avoided or, as frequently happens, an easier-to-pronounce word from the other language may be substituted in the child's active vocabulary.

Semantic development appears to follow the same general process in both monolingual and bilingual acquisition. Children go through a period of overgeneralization; for example, any animal may be referred to as *cat* or *dog*,*gatito* or *perrito*, depending upon the label first learned. Gradually, the child learns to apply the proper adult label and a cat always is a cat and never a dog. It is argued by some researchers that for bilingual children all words from both languages initially form a single vocabulary system; only gradually, as the children gain experience with the two languages, do they learn to differentiate the words of the separate languages and to use them accordingly. An additional problem for bilingual children is that the meaning of some words has different extensions in the two languages being learned. For example, the English word *brush* is used in clothes brush and paint brush whereas in Spanish a separate word is required for each kind of brush. Thus, bilingual children must learn the restrictions of the labels as they apply to corresponding items in the two languages.

Morphosyntactic structures (the order of words in sentences and forms of words such as *run* and *running*) appear to follow the same developmental order in bilingual as in monolingual children. If both languages express particular information with similar structures, those structures tend to be learned simultaneously. If a structure is more difficult in one language, it is acquired later in that language. Thus, in bilingual children, the development of certain morphosyntactic structures of one language may lag behind those of the other language because they are more complex.

**Second-Language Acquisition.** Unlike the simultaneous acquisition of two languages, second (or successive) language acquisition normally takes place after the age of 3-4 when one language—the mother tongue—has been relatively well established (but is by no means fully established). Second-language learning differs from first-language acquisition in that the learning of a second language (a) does not depend on developmental processes which are related to the developing brain chemistry, as in the case of first-language acquisition, and (b) is not the same as acquiring language skills "from scratch." Second-language learning often becomes a matter of adapting or extending existing skills and knowledge rather than learning a completely new set of skills.

The learning of a second language appears to be highly individualistic in nature. Although not much is known about how children learn a second language, their motivation and the opportunity to learn are key factors. In addition, there appear to be critical periods in the flexibility and adaptability of the brain beyond which certain aspects of second-language learning are difficult to achieve.

We do not know the optimal age for introducing a second language to children. A number of studies from various parts of the world report that children between the ages of 6 and 8 experience considerable difficulty in second-language learning. Children who begin second-language learning at the age of 9 or 10, after the first language is well established, are more successful in acquiring second-language skills than are younger immigrant or native-born children whose home language is not the dominant language of the country. However, these studies do not agree with data on Canadian immigrant children which suggest that children who arrived in Canada at older ages experienced greater educational difficulty than children who arrived prior to school entry or who were born in Canada. No definitive research on this topic has been conducted with U.S. populations. However, the informal observations of a few researchers reveal that recently arrived immigrant children from Mexico whose Spanish is firmly established are more successful in acquiring English skills than are native-born Mexican Americans.

## Assessing Language

Like other forms of educational testing, language assessment requires school psychologists to confront a number of issues and make some important decisions before engaging students in language testing. Many related issues have legal ramifications and are closely tied to the growing concern with the protection of students' basic rights. Garcia (1978) provided an enlightening analysis of the literature on social science, legislation,and litigation in relation to students' rights which must be recognized in the testing situation.

*What rights do students have with respect to:*
*the testing environment? (List at least five.)*
*the test instrument? (List at least eight.)*
*test results and their use? (List at least five.)*
*(see Garcia, 1978, pp. 456-57)*

Many present concerns with students' rights in assessment are directly related to language and culture. They include linguistically and culturally familiar test administrators; instructions and directives in a language, language level, and language style familiar to each student; test content and supporting materials that are linguistically and culturally appropriate; and tests that do not penalize the student because of language dialect differences or language limitations (e.g., when the test's language is different from the student's native or dominant language). All these factors are intended to ensure that the test results obtained for a particular student are reliable and valid. The same concerns are found in language assessment. In language testing, when the goal is to determine the language status of a student, it becomes necessary to test the student's abilities in the second or weaker language as well as in the stronger language. Nevertheless, consideration must be given to whether the student can succeed at least minimally in the second-language test. The student should not be penalized for dialect differences in either the native or second language.

*List two reasons in support of the following statement: An understanding of the nature of language and procedures for assessing language abilities is a valuable asset to the school psychologist.*

School psychologists frequently are asked to assist their school districts in preparing for and carrying out wide-scale language assessment programs. In addition, it often becomes necessary to examine a student's language characteristics prior to diagnosing learning difficulties and/or instructional planning for individual children. A thorough understanding of the nature of language and procedures for evaluating language abilities of students can be a valuable asset to school psychologists.

*In what way is language testing similar to other forms of educational testing? How does it differ? (Note similarities and differences as you proceed through the following pages.)*

The following subsection offers a discussion of some important aspects of language testing, a few of which are unique to language assessment. It is intended to assist educators in the evaluation and selection of test instruments.

### Purpose of Language Testing

What use is to be made of the information obtained? The answer to this question determines to a great extent who is to be tested and the type of instrument to be used.

Until the early 1970s, most language tests were designed for use with high school or adult students who were studying foreign languages. These tests, which still are used widely today, consist of four types:

1. **Placement:** Generally a wide-meshed screening test to determine appropriate placement in a particular sequence of courses.
2. **Achievement:** Some are criterion-referenced measures in which test items are selected to reflect specific course content. Others are norm-referenced and purport to measure a general proficiency that is considered to be typical of a particular standardization sample.
3. **Diagnostic:** A few tests are designed to pinpoint students' strengths and weaknesses in a particular language. Information from these tests is intended to assist the teacher in making decisions on instruction for individual students.
4. **Aptitude:** These tests attempt to measure language-related abilities which are thought to predict successful foreign language learning.

Since the passage of the Bilingual Education Act in 1968[1] and the subsequent legislation at the state and federal levels[2] mandating schools to provide appropriate instruction for children with limited English-speaking ability (LESA), new tests have appeared in great number on the commercial market. Much of the wide-scale language testing occurring today is directed to evaluating the language abilities of children who come from homes where a language other than English is the primary language of communication. It is this segment of the school population, rather than foreign language students, with whom school psychologists are concerned. Therefore, our discussion here focuses on procedures that are relevant to the assessment of students whose home language is other than English. It does not treat the assessment of monolingual English-speaking students nor issues related to bidialectalism.

*A bilingual student is not necessarily a fluent speaker of her/his home language.* *True*_____ *False*_____

It is generally assumed that children entering school have acquired a spoken language and that teaching them to read and write that language is a major task of the schools. Until recently, it was further assumed that in the United States the children's language would be English. However, social changes in and new immigrations into this country have forced educators to recognize that the latter assumption is unwarranted. In all regions of the country, whether rural or urban, many children enter school speaking another

---

[1] The 1968 Bilingual Education Act decreed financial assistance to local educational agencies to develop and carry out school programs which are designed to meet the special educational needs of children with limited English-speaking ability in our nation's schools (20 U.S.C. 880b, enacted January 2, 1968, P.L. 90-247, Title VII, sec. 702, 81, Stat. 816).

[2] At present, some 20 states have enacted legislation governing local educational programs for children with limited English-speaking ability. Some of the more notable court cases affecting policy at the national level are *Lau v. Nichols* (414 U.S. 563, 1974); *Serna v. Portales Municipal Schools* (499 R. 2d 1147, 10th Cir., 1974); *Aspira of New York v. Board of Education of the City of New York* (72 Civ. 4002, S.D.N.Y. August 29, 1974; also, 57 F.R.D. 62, S.D.N.Y. 1973; 65 F.R.D. 541, S.D.N.Y. 1975; 394 F. Supp. 1161, S.D.N.Y. 1975); *United States v. Texas* (Civ. No. 73-3301, W.D. Tex. 8/1/73); *Morgan v. Kerrigan* (401 R. Supp. 216, 242, D. Mass. 1975; aff'd 523 F. 2d 917, 1st Cir. 1975).

**language as well as or better than they do English. Students' linguistic skills range along a continuum from monolingualism in English, with varying degrees of bilingualism in between. Bilingual children are not homogeneous. For example, some children may have reached normal development in their native languages and be in an early stage in learning the second; others may have native-like control of both languages; yet others may have reached advanced stages in the second language but have retained only limited knowledge of their native language.**

---

*Recent research suggests that there are positive as well as negative consequences associated with bilingualism.*
*True* ______ *False*______

---

Children's school success depends largely on their ability to understand, speak, read, and write the language of the classroom. Therefore, for some children a program of instruction in a language other than English may be necessary during the period in which English language skills are being acquired. Many educators, linguists, and psychologists believe that social and academic retardation, commonly attributed to the effects of bilingualism, can be alleviated through school programs that respect the language and culture of language minority students and are sensitive to the children's emotional, social, and intellectual needs. Such programs often necessitate the use of the students' home languages in some aspects of the school curriculum for varying periods in order to ensure that concept development and academic achievement proceed normally.

What are the effects of bilingualism? The literature reveals that, for middle-income children who speak the country's dominant language, grouping classes for second-language learning so that the children do not speak the primary language of instruction ("immersion" classes) seems to be a successful way to attain high levels of functional bilingualism and academic achievement. However, for children of low socio-economic backgrounds who do not speak the dominant language, such immersion in second-language learning is not so successful. "Submersion" grouping, which mixes nonspeakers and monolingual speakers of the language of instruction, also is inadequate for them. Such groupings can lead to inadequate command of both first and second languages and poor academic achievement in general. Social, cultural, and attitudinal factors, as well as linguistic factors, are believed to be implicated in these results.

Cummins (1979) and Lambert (1978) provide excellent discussions and summaries of research on the social, psychological, and intellectual consequences of bilingualism. They summarize evidence that suggests that when the second language is viewed as an *addition* to rather than a *replacement* for the first language, access to two languages in early childhood can accelerate certain aspects of cognitive growth and can lead to high levels of competence in both languages. In contrast, in situations in which the child's home language was being gradually *replaced* by a second language, bilingualism was found to have negative effects. Many bilingual children are characterized by less than native skills in either language and inadequate cognitive and academic development.

Findings from recent studies on the effects of bilingualism have led to a current hypothesis (Cummins, 1979) that assumes the following:

1. Those aspects of bilingualism that positively influence cognitive growth probably will not come into effect until the child has reached a minimum level of competence in the second language.
2. The positive influences of bilingualism on cognitive growth probably will not occur if the child reaches only a very low level of competence in one language—either the first or the second language.
3. While offering no advantages associated with bilingualism, a minimum level of *bilingual* competence may be sufficient to avoid any negative cognitive effects.
4. A high level of bilingual competence may be necessary to lead to accelerated cognitive growth.

Legislation, at both the state and federal levels, mandates school districts to define the language status of each child and provide instructional programs commensurate with the children's language abilities. The result has been wide-scale language testing in order to classify students on the basis of language status and, subsequently, to place students into appropriate educational programs. In the United States, other forms of language testing (e.g., diagnostic, achievement, and aptitude) have received only limited attention.

### Determining Language Status

---

*Distinguish between* **language dominance** *and* **language proficiency.**

---

Language status generally is defined in terms of *language dominance* or *language proficiency*. Each is determined by a number of measures that, usually, are administered individually. They include standardized tests, standardized observations, informal appraisals, self-rating scales, and reports from others. For a discussion of specific tests see Northwest Regional Educational Laboratory (1978) and Oakland (1977). The determination of a student's language dominance or language proficiency generally is made on the basis of her/his oral language skills (understanding and speaking). However, reading and writing skills are beginning to receive greater attention, particularly with older students.

*Language dominance* refers to the relative competence a bilingual student has acquired in her/his two languages. For example, a student who is a fluent Spanish speaker and only has limited skills in English is said to be Spanish dominant. Language dominance often is determined through "indirect" measures without regard to the use of the language per se. These measures employ techniques that have virtually no formal correspondence to real-life language use. The following measures are typical:

1. **Fluency tests:** Bilingual children are asked to respond to a particular task (e.g., picture naming) in each of their two languages. Performance is scored on the basis of response time in each language.
2. **Flexibility tests:** Children are asked to give as many words or expressions as they can which are synonymous with a particular word or phrase. The test is administered in each language. The language in which a child supplies the most words or expressions is said to be the dominant language.
3. **Word Association tests:** Children are asked to respond orally to verbal or nonverbal stimuli in each of their languages. A comparison of the richness of vocabulary or clausal structure in the two languages purports to yield an indication of language dominance.
4. **Self-rating scales:** Children rate their own language abilities in each of their languages.
5. **Language Background Questionnaire:** A history of the bilingual child's past and present use of the two languages suggests his/her dominant language. This information usually is supplied by someone other than the child.

Language dominance also can be arrived at through a measure of a person's relative proficiency in two languages.

This procedure is discussed in the following subsection.

*Language proficiency* refers to the degree of competence a speaker has in a particular language. For example, adult monolingual speakers are presumed to be proficient speakers of their native language. Language proficiency usually is assessed through "direct" measures, of which standardized tests are the most frequently administered. These involve the direct and independent measurement of the person's knowledge of and ability to use each of the four skills of understanding, speaking, reading, and writing in each language in which she/he has competency. When a child's performance is compared with the performance of a normative population (usually native speakers of the language), a proficiency score is derived for each language. If the tests in each language are comparable, language dominance then can be determined on the basis of relative proficiency, that is, one language is designated as the child's dominant language by virtue of her/his greater proficiency in it.

Direct measures of language proficiency, although they provide educators with much information that may be useful in making decisions on placement and instruction, also present a number of disadvantages which must be considered.

1. They often are time consuming to administer and score.
2. They must be extensive enough to sample a wide range of skills that are included in language proficiency.
3. A speaker is not equally proficient across all the domains (e.g., topic or situations) or registers (e.g., formal vs. informal use) of language. Thus, these tests must sample a variety of language functions.
4. If the measures are to be used to determine language dominance, then they must be comparable. This is no small task when dealing with languages, especially when the languages belong to different language families.

In addition to standardized tests, direct measures such as standardized observations and informal appraisals also are used to assess language proficiency. The weaknesses of these measures are obvious. In addition to the subjectivity in assessing what *is* observed, that which is *not* observed (but may well be within the subject's linguistic repertoire) is not included in the assessment. Extensive observation (i.e., case study) by a trained observer using ethnographic techniques undoubtedly is the best way to examine a person's language performance in different communication settings and activities. However, this condition is seldom feasible in the education environment.

## Characteristics of Oral Language Proficiency Tests

### Types of Tests

*List the advantages and disadvantages of both* **discrete-item** *and* **integrative** *language tests.*

Standardized oral language proficiency tests may be categorized on the basis of item content. Most widely used tests today are either discrete-item or integrative.

*Discrete-item* tests contain a specific number of items, each of which is created to elicit a particular structure or feature of the language from the subject. Each item is scored dichotomously (right or wrong). Certain structures of subtests may be weighted and the weights incorporated into the scoring system. A proficiency level is designated on the basis of a specified range of scores.

*Integrative* tests usually are conversation based and often take the form of a structured interview. In attempting to elicit so-called "natural or free speech," they give the subjects a certain amount of freedom in structuring responses. The scoring of integrative tests may be either quantitative or qualitative in nature or it may be some combination of both. When scored quantitatively, a count or tally is made of the specified structures (e.g., noun determiners, past tense forms) in the child's speech sample. The scoring system generally includes the weighting of each structure scored. Proficiency levels are assigned on the basis of a specified range of scores.

In qualitative scoring, a subjective judgment of the child's language ability is made by a trained examiner. The child then is assigned a proficiency socre on the basis of highly defined levels. Criteria for such levels usually include (a) structural accuracy; (b) range of vocabulary; (c) range of topics; (d) degree of comprehension; (e) quality of pronunciation; and (f) degree to which errors interfere with communication.

### Modes of Testing

*Listening Comprehension is different from other modes of testing in that* ______________________________ .

A number of oral language proficiency tests include subtests that measure receptive skills (understanding) independently from production skills (speaking).

Generally referred to as listening comprehension, this mode assesses a child's ability to comprehend (decode) an adult model of a spoken utterance. The child is asked to respond nonverbally to verbal stimuli, which often requires her/him to point to or mark a picture in a series of pictures or to identify something in a picture.

Various modes have been used to assess production skills. The descriptions of some commonly used modes follow:

1. **Sentence Repetition:** Children are asked to repeat sentences that are beyond their immediate memory span. In order for such a sentence to be imitated, children must be able to comprehend (decode) the grammatical structures which are present in the model sentence and then to encode the utterance according to their own production systems, thus drawing on the semantic, morphosyntactic, and phonological rules that each knows and uses.

1. **Direct Questioning:** Children are expected to supply an answer to a direct question. The stimulus questions are calculated to elicit particular structures.

3. **Structured Response:** As in direct questioning, this mode requires children to generate language and to rely solely upon their linguistic repertoires, both for comprehension and production. Verbal stimuli are presented to children in either the form of an incomplete sentence or a direct question, followed by a tightly controlled cue. The stimuli are so constructed that they require a particular structure to be included in the response. Upon hearing the stimulus, children are expected to supply the missing information, thereby incorporating the test item into the response.

4. **Oral Composition (Tell; Retell):** Children usually are shown a picture or some other form of visual material. In the case of telling, children are asked to create an oral composition from the information in the picture. In the case of retelling, children are asked to listen to the examiner's presentation of a standard oral composition and then, immediately, to repeat the information presented in the examiner's telling.

### Content

*Most oral language proficiency tests attempt to assess and score all four subsytems of language.*
*True*______ *False*______

The four subsystems of language (phonology, morphology, syntax, and semantics) generally are considered in the construction of a comprehensive test of oral proficiency. However, certain tests purport to assess oral proficiency through the scoring of syntax alone; some focus primarily on grammatical structures and assign lesser value to phonology; others ignore phonology altogether; and some tests essentially measure just vocabulary. Most tests treat morphology to one degree or another. Very few of the present tests claim to assess more than one or two domains (usually school and/or home); even fewer purport to examine a child's ability to use different registers or styles of language.

### Psychometric Qualities

*List some of the major limitations of most currently available oral language proficiency tests.*

Students have the right to be tested with reliable and valid instruments (Garcia, 1978). Test developers have the responsibility of determining the psychometric qualities of their tests and reporting their data in technical manuals. Unfortunately, very few oral language proficiency tests in use today have demonstrated that they meet acceptable test standards. For example, some tests are standardized on a small number of children, report no suitable reliability or validity data, and fail to identify the language characteristics of children from whom the normative data were obtained. Part of the problem lies in the test developers' neglect to pursue the development and refinement of their instruments. The problem is due partly to the complexity of language and the limitations of the statistical procedures that exist today for handling the special problems of language data. Nonetheless, the use of existing measures can be justified on the grounds that (a) no better measuring device exists, (b) the state of our understanding of language assessment is partial or incomplete, (c) there is a need to make placement decisions on criteria which are not completely subjective, and (d) the need to render services and to allocate resources is urgent.

School psychologists should be aware of and counsel others on the limitations of the present language tests. Test scores alone must be used with extreme caution; coupled with observation of a student's actual language behavior in the school setting, they may contribute to a better understanding of the student's educational needs.

## Legally Mandated Assessment and Intervention

*Language assessment has become the focal point of the issues surrounding the Lau decision.*
*True______ False______*

On January 21, 1974, the U.S. Supreme Court ruled on *Lau vs. Nichols*, a class action suit. The landmark decision mandated public schools to provide for non-English-speaking children a meaningful opportunity to participate in the public educational program. The litigation had resulted from the failure of a large urban school district to provide English language instruction to a substantial number of students who did not speak English. The failure to do so was deemed to violate Section 601 of the 1964 Civil Rights Act that bans discrimination based on race, color, or national origin in programs receiving federal financial assistance. Since that decision, many inevitable questions on language assessment have been raised at both national and district levels.

Shortly after the *Lau* ruling, the Office of Civil Rights (OCR) required all districts receiving federal funds to conduct a "language survey" to identify those children whose home language was other than English. A set of guidelines was prepared and issued to all school districts. After completion of the survey, OCR compiled a list of 333 school districts which were out of compliance with the *Lau* decision. At that point language assessment became a sociopolitical and legal issue for the entire country. At its heart lay the implication that school districts found to be out of compliance with the *Lau* decision risked forfeiting federal funding for special programs. Faced with such a prospect, school officials sought guidance from OCR. In order to assist school districts to comply with the law, OCR prepared a set of recommendations which have become known as the *Lau Remedies* (U.S. Office of Civil Rights, 1975). Questions pertaining to assessment, language development, classroom placement, and program design, which normally were under the purview of the educator, linguist, psychologist, and other social scientists, became the responsibility of OCR officials. Inasmuch as the *Lau* decision was concerned with instruction for non-English speaking children, the focal point of the *Lau Remedies* became language assessment. A determination of the linguistic make-up of schools was required in order that school districts could proceed to carry out the remainder of the *Lau Remedies* that relate to program design, staffing, and so forth. Thus, it became necessary to determine each child's language status and to classify children on the basis of language. A number of states also have passed legislation to further define education requirements for their language minority populations.

### Identification of Students of Limited English-Speaking Ability (LESA)

*Classification of students as LESA was determined on the basis of language dominance rather than language proficiency. True______ False______*

The acronym LESA entered the English language shortly after OCR issued a five-level system for categorizing students' patterns of language which were used to help districts to define their populations:

A. Monolingual speaker of the language other than English (speaks the language other than English exclusively);
B. Predominantly speaks the language other than English (speaks mostly the language other than English, but speaks some English);
C. Bilingual (speaks both the language other than English and English with equal ease);
D. Predominantly speaks English (speaks mostly English but some of the language other than English); and
E. Monolingual speaker of English (speaks English exclusively). *(Lau Remedies*, 1975, p. 2)

For children falling into categories A and B, special instructional programs (English as a second language and/or bilingual education) were required. Decisions regarding instructional programs for children classified C through E remained under the purview of the local school district, except in cases where children in these categories were underachieving. For a discussion of instructional treatment and program requirements, see sections II through IX of the *Lau Remedies* (U.S. Office of Civil Rights, 1975.)

A major problem with the preceding classification scheme is its reliance on language dominance. Such a classification does not take into consideration a child's level of proficiency in either of her/his languages. Thus, a child can fall into any one of the categories and still be limited in English language skills when compared to native-speaking monolingual children of the same age.

### Identification of Limited English Proficient (LEP) Students

---

*How does the LEP classification eliminate the difficulties occasioned by the LESA classification?*

---

Recently the acronym LEP came into common use. The U.S. Department of Education requires school districts that seek ESEA Title VII funds to identify children for program placement on the basis of each student's degree of English language proficiency. Clearly, meeting this condition requires the use of some standardized measures, and test developers have been quick to respond. Most language tests now include cut-off scores to designate LEP students.

A major effort—one involving the combined effort of educators, linguists, psychologists, and statisticians—is needed to develop tests to assess the language proficiency of bilingual children and to meet the commonly accepted standards of other professions. With federal and private funds now being made available for research, it is hoped that such a major effort will take place.

### Diagnosing Learning Difficulties of Limited English-Speaking and Bilingual Children

---

*How can an understanding of bilingual children's language abilities be of value to the school psychologist?*

---

Children who speak a regional or social dialect that is different from that generally used in the school present fewer problems for the diagnostician than do bilingual children. Very often dialect differences result less in academic problems and more in social consequences (and their attendant problems). A school staff's familiarity with a range of American English dialects is an invaluable asset in dealing with children from other regions and/or social classes. It also can be useful to sensitize the classmates of these children to the reality of language variation. Speaking a dialect that differs from the language of the textbook apparently does not interfere with learning to read or comprehend a text. Teacher attitudes and expectations for such students may be a greater source of interference than the language differences (Weaver & Shonkoff, 1978). For a discussion of dialect differences see Davis (1973) and Shuy (1967).

A thorough understanding of the process of first- and second-language acquisition may be necessary for the diagnostician working with bilingual children. A thorough treatment of the acquisition process is presented in Hatch (1978) and McLaughlin (1978); for acquisition specific to Spanish/English children see Matluck (in press), Matluck and Mace (1973), and Padilla & Lindholm (1976).

The learning difficulties of bilingual children sometimes are brought to light through a comparison of the information which is obtained on verbal and nonverbal tests. However, to ensure reliable and valid results of either verbal or nonverbal tests, a child must be familiar and comfortable with the test administrator and test environment, and must understand the nature of the task she/he is expected to perform. These conditions often require the examiner to be a bilingual speaker of the child's languages and to be able to perform her/his role in a culturally sensitive manner (Matluck, 1979).

Inasmuch as limitations of language often make the diagnostician's tools unusable, it is advisable to determine a child's stages of development in both the first and second language as a possible source of the child's learning problems. This determination can be accomplished through the use of language tests that have comparable editions in each of the languages which the child speaks. Some tests yield diagnostic profiles that point up the child's strengths and weaknesses in each of her/his languages. A comparison of the child's profiles in the two languages may show that she/he can handle a particular concept (a) in both languages, (b) in neither language, or (c) in one language but not in the other. If, for example, the child can handle a particular concept (one that is congruent with her/his age level) in one language but not in the other, one can conclude that the child *knows the concept* but needs further development in the language in which she/he was unable to handle the concept. If, on the other hand, the child cannot handle the concept in either language, one should look more closely at the child's educational history, level of cognitive development, and the possibility of certain learning disorders. In any case, a thorough understanding of how the bilingual child functions in both languages can guide the school psychologist to make the necessary decisions about further testing and diagnosis.

## POSTTEST

1. In October, Juanito Garcia was referred to the school psychologist by his third-grade teacher for possible learning difficulties. The teacher reported that although Juanito continues to make grammatical errors and occasionally gropes for words, he appears to understand most of what is said to him in English and to be able to communicate with his classmates in English. However, even though he can decode most of the words accurately when reading orally in English, he retains little of what he has read. The school records indicate that Juanito entered second grade the previous year and was classified as an LEP student.

Discuss the nature of language and the language acquisition process as it relates to the above referral.

2. In the absence of a language specialist in your school district you have been asked to address a test selection committee, made up of teachers and school administrators, on the assessment of the oral language proficiency of students from non-English home backgrounds.

Prepare an outline of the major points you would want to discuss and provide a brief rationale for each.

3. Study the information included in the following case study and make tentative recommendations for further diagnosis and intervention strategies for Paulo.

### Case Study: Paulo

On February 12, Paulo's teacher submitted to the school psychologist a written request that Paulo be observed and tested for possible learning difficulties. School records and the teacher's referral revealed the following information.

**The Student:** Paulo is 9 years old; born in the Philippines; arrived in the U.S. in March of the preceding year; enrolled in the present school shortly after arrival in the U.S.; spoke no English at the time of enrollment; was assigned to a third-grade class taught by a Tagalog-speaking Filipino teacher; his class was made up of 32 children of various racial-ethnic groups, including four Ilokano-speaking Filipino children and three Tagalog speakers; his attendance during that period of enrollment was regular; attended special "pull-out" ESL classes for 30 minutes per day which were taught by a Mandarin-English bilingual teacher; he was retained in third grade the following year.

**The Parents:** Paulo presently lives with his mother, a native speaker of Ilokano, and his father, a Tagalog-Ilokano bilingual speaker. Both parents are limited in their ability to speak English. The family lives in a low socio-economic area that is served by the school. Paulo's mother works from 7:00 p.m. to 1:00 a.m. as a nurse's aide at one of the large hospitals in the center of the city; the father is employed as a custodian in a large office building and works between 7:00 a.m. and 4:30 p.m.

An interview with the parents also elicited the following information: The parents immigrated alone to the U.S. four years ago; Paulo, the youngest of four children, joined the parents in the U.S. approximately one year ago. The three older children remained in the Philippines and presently are living with the paternal grandparents; from birth until immigration to the U.S., last year, Paulo lived for varying periods of time with each set of grandparents. The maternal grandparents live in a rural area and are monolingual speakers of Ilokano; the paternal grandparents live in Manila and, although they speak some Ilokano, are predominantly speakers of Tagalog.

**Present School Situation:** Paulo reported for school on the first day of classes and was assigned to the third-grade class taught by Mrs. Guerzon, a native speaker of Tagalog but with some knowledge of Ilokano. The class consists of 30 students of various racial-ethnic groups, including four Filipino children who speak Tagalog. Paulo continues to receive ESL instruction for 30 minutes per day from his previous ESL teacher.

Mrs. Guerzon reports that Paulo is seldom successful in his classroom assignments despite his strong efforts at times. He is not able to communicate well in English and frequently shows signs of lack of understanding when spoken to in either Tagalog or Ilokano. In the last few months he has become disruptive in the classroom and extremely aggressive on the playground, provoking fights with other children.

The school nurse's records reveal that Paulo is obese and awkward in his physical movements. His eyesight and hearing appear to be normal.

**Language Assessment:** As a part of the distsrict's placement procedures for students from non-English home backgrounds, Paulo was administered comparable oral language proficiency tests in English and Ilokano in the fall of the present school year by an Ilokano-speaking ESL teacher and the school district's language specialist. The test battery was designed to assess children's ability to handle certain basic communication concepts (semantic categories), which are essential for performance in a school setting. These include the skills of identifying, classifying, quantifying, interrogating, and negating, and of showing important relations, such as spatial, case, and temporal.

A summary report of the language assessment kept in Paulo's cumulative record follows:

**Ilokano Test**

1. Paulo's composite score: 73%. (Listening Comprehension—70%; Sentence Repetition—73%; Structured Response—77%.)
2. General: His Ilokano is completely native and certainly adequate for communication and for learning tasks in that language. Regional dialect and regional usage is evident.
3. The gaps that exist indicate lack of formal schooling in Ilokano as well as lack of contact with educated Ilokano speakers. There is also quite a bit of interference from Tagalog, which is very likely due to the child's varying periods of residence in both Ilokos Sur and Manila, with much shuttling back and forth.
4. His weakest areas are negation, interrogation, and number. He communicates these concepts but his usage frequently is substandard.
5. He handles all the concepts examined except negation.

**English Test**

1. Paulo's composite score: 48%. (Listening Comprehension—63%; Sentence Repetition—39%; Structured Response—43%)
2. General: His English is in an early developmental stage in which internalization has taken place with only common vocabulary items (see "Identification") and with only a very few grammatical structures.
3. His comprehension exceeds his production abilities, which is typical of an early acquisition stage. However, indications are that he is only partially comprehending spoken English (Listening Comprehension—63%).
4. He handles all the concepts examined but is limited in his ability to verbalize them in English, with the exception of English negation. Areas of greatest weakness are grammatical number, interrogation, spatial relations, classification, and case relations.

**Comments**

1. Established norms for this test do not exist but a comparison of his scores on the two tests is quite revealing, as indicated by the preceding analyses.
2. Paulo is dominant in Ilokano. *The areas of number and interrogation show decided weakness in both languages.*
3. Paulo's taped responses indicate the need for special instruction in both languages.
4. We strongly recommend, as a minimum, special oral English instruction on a regular basis in the following areas:

**Morphology**

inflectional endings /-s/, /-z/ in noun plurals, possessive, and subject-verb agreement.
(b) inflectional forms of the verb to show temporal relations: /-t/, /-d/, /-id/.
(c) pronouns: *he, she, it* (gender indicators), forms and usage for all case relations.
(d) adjectivals; formation and placement of possessive adjectives, descriptive adjectives, and adverbials.

**Syntax and Vocabulary**

(a) selection and inclusion of prepositions and articles.

(b) English patterns of interrogation.

4. Your school district is interested in submitting a proposal to the U.S. Department of Education to secure ESEA Title VII funds to initiate an instructional program to meet the needs of children with limited English-speaking ability. You, along with the district's language specialist, have been asked by the school administration to select a language assessment instrument for use in the classification and placement of students.

List the general requirements of the instrument you would select and justify the use of such an instrument.

## SIMULATION 1

Type: Case Study of "Sandra"

Purpose: To improve the school psychologist's skill in analyzing and interpreting the effects of language on the school behavior of bilingual children.

**Step I**

**Materials:** Script of background information.

**Activity:** Based on the background information, you should develop a plan of action for diagnosing Sandra. The model response for this activity is included in Step II.

### Background Information

Sandra's teacher submitted a written request in October that the school psychologist observe and test Sandra for possible learning difficulties. School records and the teacher's referral revealed the following information:

**The Student:** Sandra is 8 years old; born in Mexico; arrived in the U.S. in April of the preceding year; enrolled initially in the present school shortly after arriving in the U.S.; spoke no English at the time of enrollment; was assigned to a first-grade class taught by a monolingual English-speaking teacher; her class contained 28 monolingual speaker of English; Sandra's attendance during the period of her enrollment (the last six weeks of school) was intermittent; she was retained in first grade for the following year.

**The Parents:** Sandra lives with her mother and stepfather (non-Hispanic surname) in a low socio-economic area served by the school; her mother attends Far West Beauty School from 10:00 a.m. to 6:00 p.m.; the school records contain a phone number for a babysitter, Mrs. A. E. Swanson, who lives next door to the family. No information is available from the school records on the stepfather or other family members.

**Present School Situation:** Sandra was brought to school by her mother on September 15, four days after school started; she was assigned to her present class that contains 26 monolingual English-speaking first-grade children. The class is taught by Mrs. Smith, a monolingual English speaker.

Mrs. Smith reports that Sandra is not learning to read, refuses to do her work, is sullen when spoken to by the teacher, seldom responds orally when spoken to by either the teacher or the children, causes disruptions in the room, and fights with children on the playground.

The school nurse's records show no signs of sight or hearing problems nor physical abnormalities.

A survey of the school staff reveals no speakers of Spanish among the professional or support personnel.

**Step II**

**Materials:** A plan for diagnosing and prescribing appropriate instructional treatment for Sandra.

**Activity:** Discuss and evaluate with another psychologist the following plan of action for diagnosing and prescribing instructional treatment for Sandra.

### A Plan of Action

1. Observe the child in different school settings.
2. Investigate the school district's resources for determining the language characteristics of the child.
3. Arrange for in-depth language assessment of Sandra's skills in both English and Spanish.
4. Interview the parents and obtain a case history of Sandra's previous educational history.
5. Analyze and summarize all information. Discuss findings with the teacher and other relevant school personnel.
6. Prepare recommendations and write instructional prescription.

**Step III**

**Materials:** Script of additional information.

**Activity:** After considering all accumulated information, prepare recomendations and instructional prescription for Sandra.

(The model response for this activity is included in Step IV.)

### Additional Information

The following information was gleaned from various sources:

**Observations:** Sandra is physically larger than most of the children in her first-grade class; her present table and chair are too small; she is "undisciplined" in the room (she frequently stands and walks around the room and interferes with other children at work and annoys both the teacher and her classmates); she seems happiest when working at the "listening center," which contains audio tapes, picture books, etc. On the playground she is aggressive; she appears not to understand or to ignore "turn-taking" amenities; frequently isolates herself from the group when on the playground or in the lunchroom; walks home alone and sometimes crosses the street to avoid contact with children walking in the same direction.

**Language Assessment:** A series of calls to the central school administration offices revealed that the district's resources included a trained staff of bilingual English-as-a-Second-Language (ESL) teachers who were assigned to schools with large numbers of limited-English-speaking students. In addition, the resources included the services of a consultant linguist—himself a fluent speaker of Spanish—who was brought into the district to develop oral proficiency tests for use with the language groups in the district. A formal request was submitted to the district for an examination of Sandra's language abilities in English and Spanish. Following a series of discussions among the school psychologist, Sandra's teacher, the linguist, and one of the ESL teachers, a battery of comparable oral language proficiency tests in Spanish and English was administered to Sandra by the linguist and ESL teacher. The test battery was designed to assess children's ability to handle certain basic communication concepts (semantic categories) which are essential for performance in a school setting. They include the skills of identifying, classifying, quantifying, interrogating, and negating, and of showing important relations, such as spatial, case, and temporal.

Within two weeks the school psychologist received a copy of the scored protocols and a summary report that included the following information:

### Spanish Test

1. Her composite score: 78%. (Listening Comprehension—74%, Sentence Repetition—87%, Structured Response—73%.)
2. General: Her Spanish is completely native and is adequate for communicating and learning tasks in that language. Regional dialect and usage are evident.
3. The few gaps that exist suggest lack of formal schooling in that language as well as lack of contact with educated Spanish speakers.
4. Her weakest areas are spatial relations, interrogation, and negation. She communicates these concepts, but her usage frequently is nonstandard.
5. She handles all the concepts examined.

### English Test

1. Her composite score: 52%. (Listening Comprehension—78%, Sentence Repetition—37%, Structured Response—40%.)
2. In general, her English is in an early developmental stage in which she has internalized only the most common vocabulary items and a few grammatical structures.
3. Her comprehension is far ahead of her production abilities (typical of an early acquisition stage).
4. She handles all the concepts examined but is limited in her ability to verbalize them in English.

**Interview with Mother:** The school psychologist arranged for the mother to meet with the linguist and the ESL teacher to gather information on Sandra's previous educational experience and her current behavior at home. Because the mother's ability to communicate in English was somewhat limited, the interview was conducted in Spanish. The interview revealed the following information:

1. Prior to arriving in the U.S. at age 7½, Sandra had lived with an elderly grandmother in a small village in Mexico. When she was 2 years old, the mother moved to Mexico City, worked there for a few years, and emigrated to the U.S. following marriage to her present husband—an Anglo U.S. citizen. Prior to Sandra's arrival in the U.S., contact between the mother and child had been occasional and of short duration because the child had remained with the grandmother from the age of 2 on.

2. At the age of 6, Sandra attended one month of school in her Mexican village. She had been unhappy, refused to conform to the school's requirements, and was sent home by the school officials. She had no further schooling in Mexico. Her day was filled, for the most part, with household chores and helping her grandmother to earn a living by vending food in a market.

3. The mother reported that Sandra has not adjusted well to living in her home in the U.S.; she is described as rebellious, destructive with household items, untidy with her clothes and belongings, and difficult for both mother and stepfather to manage; she frequently loses things, such as lunch money and belongings; does not want to go to school; and frequently complains of head and/or stomach pains. Sandra's medical checkup at a local clinic revealed no physical abnormalities.

### Step IV

**Materials:** List of recommendations and instructional prescription.

**Activity:** Discuss and evaluate the following list of recommendations and instructional prescriptions as they relate to Sandra.

## Recommendations and Instructional Prescription

### Recommendations

1. For the present, retain Sandra in the regular school program with frequent and systematic observations to be carried out by the school psychologist in collaboration with the classroom teacher.
2. Reassign Sandra to a third-grade classroom where she will be with children her own age.
3. Counsel with the third-grade teacher on Sandra's past history and the prescribed instructional treatment.
4. Arrange for one of the district's ESL teachers to consult and to provide assistance to the third-grade teacher as needed.
5. Arrange to have Sandra participate in after-school, noninstructional activities (e.g., Bluebirds, Girl Scouts, park and community recreational activities).
6. Investigate the educational resources in the school district for possible bilingual instruction for Sandra.
7. Reevaluate recommendations in early December of this year.

### Instructional Prescription

1. Provide special instruction in oral English on a regular basis.
2. Eliminate instruction in reading in English for the present. Substitute individualized listening and prereading activities at the listening center.
3. Ensure Sandra's full participation in subject areas where language is not so crucial (e.g., P.E., Art, Music). This will undoubtedly require enlisting the cooperation of her classmates in assisting her in understanding directions and procedures in such activities.
4. Establish an appropriate and interesting routine for Sandra in the classroom and gently, but firmly, insist on her compliance.
5. Encourage a positive self-concept through providing activities in which Sandra can succeed. Provide positive feedback when appropriate.

## SIMULATION 2

Type: In Basket.

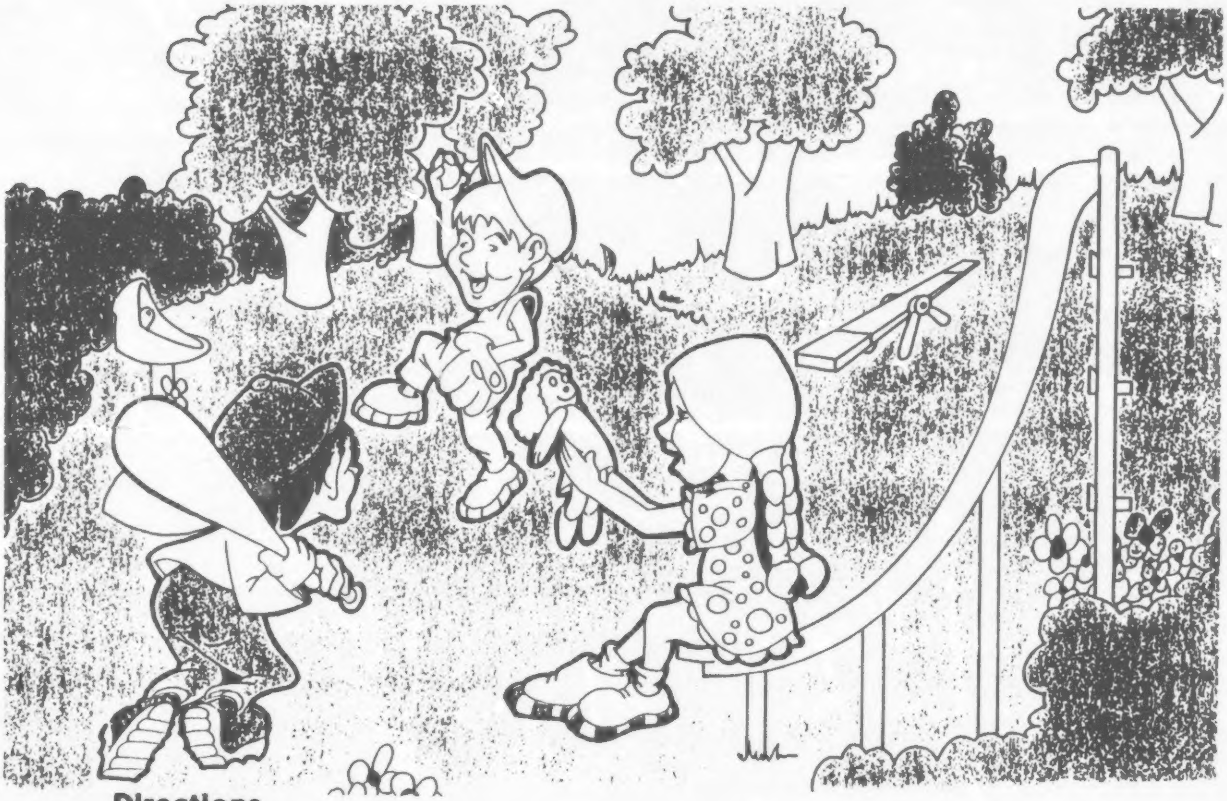
Purpose: To improve the school psychologist's skill in evaluating the usefulness of oral language tests for assessing the oral language proficiency of young bilingual children.

**Materials:** A completed evaluation checklist on an oral proficiency test and sample items and visual stimuli from the test materials.

**Activity:** On the basis of the information provided and sample items and visual material, evaluate the usefulness of the instrument for assessing the oral language proficiency of young Spanish-English bilingual children.

The following information was acquired on a popular test. Consider this information and the following test in arriving at your evaluation.

| | |
|---|---|
| 1. Language of the test | Both Spanish and English editions |
| 2. Test materials | Examiner's manual; technical manual (2 pages), consumable student booklet; stimulus pictures |
| 3. Cost of materials | $6.00 per class |
| 4. Purpose of the test | "This instrument measures and compares a student's oral proficiency in English and Spanish." (p. 1) |
| 5. Type of test | Individually administered; discrete items. Questions are asked to elicit oral responses based on pictures. |
| 6. Content | "Questions are asked to elicit responses that contain a comprehensive range of syntactical and morphological elements normally found in the language repertoire of five and six year old native speakers of the language." (p. 1) |
| 7. Age range recommended | Grades K-3. A second-level test is available for grades 4-6. |
| 8. Normal vs. unusual language use | Specific vocabulary and context use. Natural only to instructional setting. Uses only one test mode (question-answer). |
| 9. Reliability | |
| A. Internal consistency | None reported in published materials. Scoring chart contains some inaccuracies (p. 6). Test cannot produce certain of the scores on the chart. |
| B. Test-retest | None reported in published materials. |
| C. Interrater | None reported in published materials. Manual provides criteria and some examples for judging acceptability of item response. |
| 10. Validity | |
| A. Content | None reported in published materials. It appears to cover some of the high frequency lexical and grammatical structures. However, the stimulus sentences are not structured tightly enough to assure the elicitation of intended structures. |
| B. Criterion-related | None reported in published materials. |
| C. Construct | None reported in published materials. However, the nature of the test implies a certain degree of construct validity. |
| 11. Item quality | Loosely structured; scoring ambiguities in what is acceptable; no item analyses reported. |
| 12. Visual quality | Clear and attractive but perspective is lacking in some instances. |
| 13. Audio quality | No audio materials are required. |
| 14. Norming data | Field tested with a sample population that included primary students from a predominantly monolingual English-speaking school and from a predominantly Spanish-dominant school. The two schools represent the extremes in language and socio-economic differences found in an urban school district located in close proximity to the U.S.-Mexican border. The size of the field-test population (number) is not stated nor are frequency distributions given. |
| 15. Programmatic response | Programmatic response is suggested (pp. 9-10). Information is interpretable by a teacher, but is limited in its brevity. |
| 16. Administration time | Manual states "approximately seven minutes for each language" (p. 2). Reviewer feels that 7-15 minutes per language is realistic. |
| 17. Training of examiners | No specific materials or training sessions are required by the test developer or publisher. |



## Directions

The Oral Language Dominance Measure is administered in a conversational and informal manner. Have an informal conversation with the student. Rapport must be established. Under no circumstances should the student be told that this is a test. Tell the student you will be showing him/her some cartoon pictures and asking some questions. You may use the picture booklet to make conversation. However, do not ask any of the test questions. Begin the test once the student is comfortable and relaxed. Allow the student to respond in a conversational manner. Do not ask him/her to answer in complete sentences. Accept any response the student gives.

**Show the student picture 1.**

Ask the student to point to the following

**(1) ball (2) bat (3) doll (4) glove**

**If the child cannot complete this task,** show him/her the items asked for. This is an activity to acertain if the student understands what each item in the picture represents. Responses are not to be recorded

1. Point to the play area.
**WHERE ARE THEY PLAYING?** ____________________

2. Point to the children.
**WHO'S PLAYING THERE?/**
**WHO IS PLAYING THERE?** ____________________

3. Point to the doll
**WHOSE DOLL IS IT?** ____________________

4. Point to the girl
**WHAT'S SHE DOING?**
**WHAT IS SHE DOING?** ____________________

**Do not continue** if the student has not been able to respond to three of the four questions. Indicate reason for not responding.

☐ Shy ☐ Ill ☐ Lack of Language ☐ Other

Explain ____________________

____________________

Continue the **Measure** if the student has answered three of the four questions. If the student fails to answer three consecutive items in any language, stop testing

### Simulation 2: Response Model

From a strictly practical point of view, this test offers several distinct advantages: It requires little time for administration and to train examiners, its cost is relatively inexpensive, forms exist in both Spanish and English, and the test offers suggestions for interventions.

The test also has a number of limitations as an oral language proficiency measure for young bilingual children. Among the most serious are the lack of data on the psychometric qualities of the test and on the standardization sample; ambiguities in the scoring procedures, which will most assuredly affect reliability of the measure; content validity of the test may be questioned in that the relatively small number of items appear to be limited to high-frequency lexical and grammatical structures; and the single mode of testing used by the instrument may tap only language usage typical of an instructional setting and may not assess a representative sample of the child's over-all language proficiency.

This test probably yields information on dominant language but slights information needed for classifying and placing students, planning instructional programs for groups of children, and diagnosing learning difficulties and developing intervention strategies for individual children.

## PRETEST KEY

| Item Number | Correct Response | Item Number | Correct Response |
|---|---|---|---|
| 1 | b | 21 | c |
| 2 | c | 22 | a |
| 3 | c | 23 | d |
| 4 | a | 24 | b |
| 5 | c | 25 | d |
| 6 | d | 26 | a |
| 7 | a | 27 | c |
| 8 | c | 28 | d |
| 9 | c | 29 | b |
| 10 | a | 30 | b |
| 11 | d | 31 | c |
| 12 | c | 32 | d |
| 13 | a | 33 | b |
| 14 | a | 34 | d |
| 15 | c | 35 | b |
| 16 | d | 36 | b |
| 17 | b | 37 | b |
| 18 | c | 38 | a |
| 19 | b | | |
| 20 | d | | |

## REFERENCES

Cummins, J. Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 1979, 49(2), 222-51.

Dale, P. S. *Language development: Structure and function.* Hinsdale, Ill.: Dryden, 1972.

Davis, A. L. (Ed.). *A resource book for teachers: Culture, class, and language variety* (2nd ed.). Urbana, Ill.: National Council of Teachers of English, 1973.

English Language Services, Inc. English pronunciation: A manual for teachers. *Collier-MacMillan Teacher's Library.* New York: MacMillan, 1968.

Fromkin, V., & Rodman, R. *An introduction to language.* New York: Holt, Rinehart & Winston, 1974.

Garcia, G. N. Test encounters of the third kind. In J. E. Alatis (Ed.), *Georgetown University round table on languages and linguistics 1978: International dimensions of bilingual education.* Washington, D.C.: Georgetown University Press, 1978.

Hatch, E. M. (Ed.). *Second language acquisition: A book of readings.* Rowley, Mass.: Newbury House, 1978.

Kehoe, M. (Ed.). Applied linguistics: A survey for language teachers. *Collier-MacMillan Teacher's Library.* New York: MacMillan, 1971.

Lado, R. *Linguistics across cultures: Applied linguistics for language teachers.* Ann Arbor: University of Michigan Press, 1957.

Lambert, W. E. Some cognitive and sociocultural consequences of being bilingual. In J. E. Alatis (Ed.), *Georgetown University round table on languages and linguistics 1978: International dimensions of bilingual education.* Washington, D.C.: Georgetown University Press, 1978.

Langacker, R. W. *Language and its structure: Some fundamental linguistic concepts.* New York: Harcourt, Brace and Jovanovich, 1968.

*Lau vs. Nichols.* 414 U.S. 563-572 (January 21, 1974).

Matluck, J. H. Bilingualism of Mexican-American children: Language characteristics. In E. L. Blansitt, Jr., & R. V. Teschner (Eds.), *A Festchrift for Jacob Ornstein: Studies in General and Sociolinguistics.* Rowley, Mass.: Newbury House, in press.

Matluck, J. H. Cultural norms and classroom discourse: Communication problems in the multiethnic school setting. *The Modern Language Journal*, 1979, 63(4), 187-92.

Matluck, J. H., & Mace, B. J. Language characteristics of Mexican-American children: Implications for assessment. In B. N. Phillips & T. Oakland (Eds.), *Assessing minority group children: A special issue of Journal of School*

*Psychology*. New York: Behavioral Publications, 1973.

McLaughlin, B. *Second-language acquisition in childhood*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978.

Moulton, W. G. *A linguistic guide to language learning* (2nd ed.), New York: Modern Language Association, 1970.

Northwest Regional Educational Laboratory. *Assessment instruments in bilingual education: A descriptive catalog of 342 oral and written tests*. Los Angeles: National Dissemination and Assessment Center, 1978.

Oakland, T. (Ed.). *Psychological and educational assessment of minority children*. New York: Brunner/Mazel, 1977.

Saville-Troike, M. *Foundations for teaching English as a second language: Theory and method for multicultural education*. Englewood Cliffs, N.J.: Prentice-Hall, 1976.

Shuy, R. W. *Discovering American dialects*. Urbana, Ill.: National Council of Teachers of English, 1967.

U.S. Office of Civil Rights. Task force specifying remedies available for eliminating past educational practices ruled unlawful under *Lau vs. Nichols*. Mimeo memo to state and local education agencies, August 1975.

Wilkins, D. A. *Linguistics in language teaching* (2nd ed.). Cambridge, Mass.: The MIT Press, 1975.

## ANNOTATED BIBLIOGRAPHY

J. Cummins. Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 1979, 49(2), 222-251.

This article undoubtedly is one of the most significant to appear in recent years on the effects of bilingualism on cognitive growth. Its central focus is on the positive and negative effects of bilingualism on cognition, which may be related to the level of competence a bilingual reaches in each of her/his languages. Two hypotheses are formulated and combined to arrive at the position that (a) the development of competence in a second language is partially a function of the type of competence already developed in the first language when extensive exposure to the second language begins, and (b) there may be threshold levels of linguistic competence which a bilingual child must reach in order to avoid negative cognitive consequences and also to allow the potentially beneficial aspects of bilingualism to influence her/his cognitive and academic functioning.

English Language Services, Inc. English pronunciation: A manual for teachers. *Collier-MacMillan Teacher's Library*. New York: MacMillan, 1968.

This small volume is one of two short, introductory manuals designed for teachers who have little formal training in linguistics but are interested in language or language teaching. Among the topics treated are teaching and learning the sound system of a language; features and patterns of sounds and symbols; speech organs and their functions; descriptive analysis of the sound system of a language; and the phonology of American English.

M. Kehoe (Ed.). Applied linguistics: A survey for language teachers. *Collier-MacMillan Teacher's Library*. New York: MacMillan, 1971.

One of a series of two short introductory manuals, this book is designed for teachers who have little formal training in linguistics but are interested in language or language teaching. Included are discussions of the nature of language, language learning, language teaching approaches, teaching English abroad, and teaching English as a Second Language in English-speaking areas.

J. H. Matluck. Cultural norms and classroom discourse: Communication problems in the multiethnic school setting. *The Modern Language Journal*, 1979, 63(4), 187-92.

This article discusses the three sets of behavior and discourse rules that are in operation in a school setting: those of the child, teacher, and school, each interacting with the other two. The communication problems which can occur in teaching and learning as a result of the interaction are magnified and made more complicated in the case of a multiethnic classroom. Reference is made in the article to works treating the well-known sociological, psychological, economic, and political factors that affect communication in this teaching-learning situation. Beyond these, other critical factors, which are purely crosscultural and linguistic in nature, are identified and illustrated with specific examples from school settings containing a large percentage of Chinese, Filipino, and Hispanic populations.

J. H. Matluck & B. J. Mace. Language characteristics of Mexican-American children: Implications for assessment. In T. Oakland & B. N. Phillips (Eds.), *Assessing minority group children: A special issue of Journal of School Psychology*. New York: Behavioral Publications, 1973.

This publication is an in-depth treatment of the language features that characterize the speech of 5 and 6-year-old Mexican American children, the host of influences exerted on each language, and the relation of these features to how well the children function in the two language settings in which they live. Assessment considerations are suggested to evaluate more accurately the actual abilities of the children and to provide more specific bases for planning improved educational programs for them.

B. McLaughlin. *Second-language acquisition in childhood*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978.

A very readable, comprehensive overview of research findings on both first- and second-language acquisition, the book also contains a discussion of the effects of bilingualism on the individual and an evaluation of widely used second-language teaching methodologies and program types.

T. Oakland (Ed.). *Psychological and educational assessment of minority children*. New York: Brunner/Mazel, 1977.

One of the most complete sourcebooks available, this volume offers an effective and relevant assessment system in which diagnostic processes are joined to viable intervention strategies. Among the topics treated are professional, legal, social, and ethical issues. Of special value are the appendices; they contain relevant portions of ethical and legal documents and an annotated bibliography of a number of language tests.

M. Saville-Troike. *Foundations for teaching English as a second language: Theory and method for multicultural education*. Englewood Cliffs, N.J.: Prentice-Hall, 1976.

This is an indispensable text for ESL teachers. It is "basic training" for anyone involved in the school experiences of speakers of languages other than English and contains the necessary background information in the areas of language learning, linguistics, and culture. It also discusses strategies for instruction and the role of ESL in bilingual education.

# NONBIASED EDUCATIONAL ASSESSMENT

Alan M. Hofmeister
and
Caroline N. Preston
*Utah State University*

## Table of Contents

Objectives ........ 77

Pretest ........ 77

The Testing Debate ........ 78
    Key Conceptual, Methodological, and Ethical Assessment Issues ........ 78
    Clarification of the Issues ........ 78

The Educational Role of Professionals and PL 94-142
    The School Psychologist and Educational Assessment ........ 79
    PL 94-142 and Nondiscriminatory Evaluation ........ 79

Assessment Methodology ........ 79

Relevant Assessment Variables ........ 79

A Model for Systematic Assessment ........ 80
    Screening ........ 80
    Evaluation of Screening Data and Planning Interventions ........ 81
    Intervention ........ 81

Norm-Referenced and Criterion-Referenced Assessment ........ 81

The Process of Diagnosis ........ 83
    Diagnostic Devices ........ 83
    The Relationship of Domain-Referenced Testing to Criterion-Referenced Tests ........ 84
    Adapting Test Content ........ 86
    Random and Linear Access ........ 86
    Cross-Referencing ........ 86
    Commercial Criterion-Referenced Testing Systems ........ 86
    Setting a Criterion ........ 86

Use of Profiles in Planning Educational Programs ........ 87

Evaluation of Intervention Effects ........ 88

Interpretation and Communication of Assessment Information ........ 88
    Administrators ........ 88
    Teachers ........ 88
    Parents ........ 89
    Students ........ 89

Conclusion ........ 89

Posttest ........ 89

Simulation Exercises ........ 91

Key to Pretest ........ 93

Key to Posttest ........ 94

References ........ 95

Recommended Readings ........ 95

## OBJECTIVES

On completing this module, participants will be able to identify appropriate and inappropriate procedures for assessing students with consideration of the following conceptual, methodological, and ethical issues:

1. The testing debate.
2. The educational role of professionals and P.L. 94-142.
3. Assessment methodology.
4. Relevant assessment variables.
5. Assessment procedure.
6. Norm-referenced and criterion-referenced assessment.
7. The relationship between assessment and instruction.
8. Evaluation of intervention effects.
9. Interpretation and communication of assessment information.

## PRETEST

The following items are intended to assess your understanding of the significant issues in educational assessment.

**Short Answer Items**

1. Define the purpose of educational assessment.
2. What is the difference between testing and assessment?
3. Outline the requirements of PL 94-142 as they relate to the educational assessment of minority students.
4. Identify two sources of information about a student's functioning in educational settings.
5. Identify two factors to be considered when making informed decisions about a student's educational functioning.
6. Identify the major components of a systematic assessment procedure.
7. Identify two sources of inconsistency or error in methods of gathering assessment information about a student.
8. Describe the nature of the information that can be gathered from achievement tests.
9. Identify two factors that can negatively influence the validity of a test.
10. What are the two purposes of educational diagnosis?
11. What is the difference in the nature of information obtained from a norm-referenced test and a criterion-referenced test?
12. Describe domain-referenced testing.
13. What are the five major considerations in setting a criterion for a student's performance?
14. State the two areas which are of concern when evaluating treatment effectiveness.
15. Define treatment validity.
16. What are the three mandates laid down in PL 94-142 that relate to the interpretation and communication of assessment information?

In front of the next 11 items, place a "T" if you think the statement is true, or an "F" if you think the statement is false.

_____ 1. Tests are either "good" or "bad." Poor programming is the result of using "bad" tests.

_____ 2. The purpose of assessment is to prescribe a fixed treatment that continues until a cure is achieved.

_____ 3. The requirements of PL 94-142 necessitate a comprehensive approach to assessment team membership and to instrumentation.

_____ 4. School psychologists can be most useful by confining their activities to the administration of standardized tests of ability and achievement.

_____ 5. When placement decisions for a student are made, information from norm-referenced tests provides the most comprehensive data-base.

_____ 6. Information obtained from a student's parents may assist in verifying information given by the referring teacher.

_____ 7. Students' achievement test scores reflect both their ability and the instructional effectiveness of their educational program.

_____ 8. Human error is a minor source of inconsistency in the assessment process and compensations for unreliability can easily be made.

_____ 9. It is the responsibility of the test developer to ensure test validity. Once this has been demonstrated, the professional need have no concern about using the test.

_____ 10. An advantage in using norm-referenced tests is their adaptability to a specific curriculum sequence.

_____ 11. To be of maximum use to the professional, criterion-referenced testing must be integrated into the day-by-day functioning of the classroom and must not be separated out as a "testing" activity.

## The Testing Debate

The role of testing within the educational system has long been a matter for debate. A view of testing as a sorting procedure which results in students being divided into those who will succeed and those who will not is common. Testing is sometimes considered a tool for social decision-making about the futures of millions of children (Burton, 1978). Such decision-making, in too many cases, may result in inappropriate placement of children in educational programs and curricula, and failure in postschool experiences. It has been noted that inappropriate decisions seem to occur more frequently for certain minority groups and, to many, this has seemed discriminatory or unfair.

In the testing debate, antagonists cite not only political and social abuses of testing (Kamin, 1975), but also testing devices that are technically inadequate, are normed on populations which bear no resemblance to the pupils being tested, or inadequately sample the behaviors that the instrument supposedly tests (Salvia and Ysseldyke, 1978). Adelman and Taylor (1979) present several conceptual, methodological and ethical assessment issues which merit discussion.

### Key Conceptual, Methodological, and Ethical Assessment Issues

A. Conceptual

Current practices reflect a rather inadequate appreciation of the following matters:

1. Whose interests are and should be served by the assessment activity (e.g., the society, the client, the intervener)?
2. Who sets criteria determining assessee status, needs, problems, and progress?
3. Who decides on whose interests should be served and on who is to set criteria?
4. What are the various tasks which currently are subsumed under the term "assessment diagnosis"?
5. What are the limitations in perspective of the models of cause and correction upon which current activity is based?
6. What are the assessor biasing factors which need to be systematically accounted for in assessment activity?
7. What are the assessee motivational and developmental factors which need to be systematically accounted for in assessment activity?

B. Methodological (technical, practical)

While widely acknowledged, the following problems continue to place major limitations on assessment activity:

1. The more complex the assessment objectives, the lower the reliability of the total set of measurement procedures tend to be.
2. Construct validity often has not been demonstrated satisfactorily by scientific standards, e.g., the "validation" procedures, when undertaken, often are tautological.
3. Predictive (and postdictive) validity appears to diminish, in some instances at an exponential rate, the more distant in time the assessment data being gathered are from the criterion being predicted to (or from the original factor causing the behavior under investigation).
4. There is a sparsity of systematically gathered and agreed upon norms and standards for making judgments ("good-bad," "normal-abnormal," "success-failure") thereby resulting in idiosyncratic variations on judgments which are often beyond accountability.
5. The utility of a procedure may be judged as much (or more) on the basis of its marketability and the current absence of a feasible alternative, as on its efficacy (e.g., its validity with regard to the decisions being made, its ability to add information beyond base rate levels).
6. The costs of assessment practices are escalating; time demands often are extensive; referral practices tend to overrely on "old boy" networks, etc.

C. Ethical

Ethical practices reflect inadequate efforts to systematically detect and minimize the following iatrogenic effects:

1. Misidentification of the cause of a problem (e.g., false negatives and postives such as those resulting from over-reliance on person-focused tests and observer-assessor biases).
2. Misprescriptions related to subsequent intervention procedures.
3. Violations of rights (e.g., failure to get truly informed consent, invasion of privacy, denial of access to assessment reports and of the right to correct the record).
4. Negative repercussions of assessment processes or products (e.g., increasing feelings of anxiety, incompetency, and lack of self-determination, increasing overreliance and dependency on professionals, initiating self-fulfilling prophecies and stigmatizing effects).
5. Failure to pursue professional responsibility with regard to improving standards of practice and advancing knowledge (including collusion with an inadequate status quo).

—from Adelman, H. S., & Taylor, L. Initial psychoeducational assessment and related consultation. *Learning Disability Quarterly*, 1979, 2, 52-64.

**Clarification of the Issues**

Consideration of the following three points may help clarify these testing issues:

1. Concentration on the abuses that have occurred in psychoeducational assessment may blur the distinction between tests and the decisions made from test data. Many abuses occur because of ignorance or overzealousness on the part of the diagnostician or decision-maker. Tests are not necessarily "good" or "bad"; such value judgments can be reserved for evaluating test use in screening, placement, program planning, and evaluation of individual pupil progress.

2. Testing is but one method of assessing a child and is only one tool to use in making educational decisions. The diagnostician who fails to make use of such techniques as the interview, observation, and environmental analysis, as well as his/her own professional judgment has inadequately gathered information to make educational decisions. (Shertzer and Linden, 1979)

3. "The process of assessment is a constructive . . . flexible, continuous process, leading not to a fixed prescription of treatment until a cure is achieved, but to an ongoing program which may frequently be modified in the interests of the student's life situation and of a reduction in his current specific difficulties." (Clarke and Clarke, 1975)

## The Educational Role of Professionals and PL 94-142

Individual roles and responsibilities of professionals in educational settings are, of course, determined by the nature of their professional training. The result has often been a *compartmentalization of responsibilities* and a view of the child as a group of separate and unique entities. This is no longer feasible under Public Law 94-142, which requires that a student's individual education program (IEP) be "developed by a team consisting of the child's teacher, a person other than a teacher who is qualified to provide or supervise the provision of special education, one or both parents, the child (when appropriate), and other persons who are brought in at the discretion of the school." The purpose of the IEP is to provide an overall program of special education and related services and it should include a justification for those services and placement. In addition, the IEP must include the objective criteria that will be used to evaluate the child's achievement. A multidisciplinary assessment team and multiple assessment methods are necessary.

### The School Psychologist and Educational Assessment

The traditional role of the school psychologist often restricts him/her to the administration of standardized tests of ability and achievement. Information from these tests is used to *predict* a student's success or failure during school and in later life. Diagnosis of a student's abilities or deficits may be confined to norm-referenced statements based on composite scores and often has few direct implications for instructional intervention. Such diagnostic practices were enlisted to support special education practices of classification, according to handicap for the purposes of federal funding.

Reynolds (1975) calls attention to changes in special education that are guiding changes in assessment practices. He states:

> We are in a zero-demission era; consequently, schools require a decision orientation other than simple prediction; they need one that is oriented to individual rather than institutional payoff. In today's context the measurement technologies ought to become integral parts of instruction designed to make a *difference* in the lives of children and not just a *prediction about* their lives. (p. 15)

### PL 94-142 and Nondiscriminatory Evaluation

In order to ensure appropriate educational placement, equality of educational opportunity, and the right to ethnic dignity and respect, and to prevent unfair stigmatizing of students, the Education for All Handicapped Children Act of 1975 (PL 94-142) mandates the following assessment procedures:

> 1. The testing and evaluation materials and procedures will be selected and administered so as not to be culturally discriminatory.
> 2. Such materials and procedures are to be provided in the child's native language or mode of communication.
> 3. No single procedure or test can be the sole criterion for determining the appropriate educational program for the child (Section 612).

Included in the Act, the following recommendation is made to education agencies:

> A procedure also should be included in terms of a move toward the development of diagnostic-prescriptive techniques to be utilized when for reasons of language differences or deficiencies, non-adaptive behavior, or extreme cultural differences, a child cannot be evaluated by the instrumentation of tests. Such procedures should insure that no assessment will be attempted when a child is unable to respond to the tasks or behavior required by a test because of linguistic or cultural differences unless culturally and linguistically appropriate measures are administered by qualified persons. In those cases in which appropriate measures and/or qualified persons are not available, diagnostic-prescriptive educational programs should be used until the child has acquired sufficient familiarity with the language and culture of the school for more formal assessment. These evaluation procedures should also assure that persons interpreting assessment information and making educational decisions are qualified to administer the various measures and qualified to take cultural differences into account in interpreting the meaning of multiple sets of data from both the home and the schools. (p. 29)

---

### *Task 1*

*a. School Psychologists often encounter the statement that they are "test-happy," and that the tests they give are of little help to the classroom teacher. What is the philosophy regarding psychological testing in your school district? To what extent are you closely involved with classroom teachers? Can your relationship with them be described as cooperative?*

*b. Review the conceptual, methodological, and ethical issues and problems related to assessment which appear at the beginning of this section. Use the questions and statements there to evaluate professional practice in your school district.*

*c. Review the PL 94-142 recommendation to educational agencies. To what extent are you familiar with the nature of the cultural differences among children with whom you interact? How familiar are you with the cultural patterns of these children and to what extent have you sought information about expectations and demands that their society (both in school and at home) makes on them?*

---

## Assessment Methodology

Educational assessment is a multifaceted and systematic process carried out for the purpose of making decisions about the performance of students in their current ecology. Assessment is more than testing. It is the systematic process of using information from all possible sources in order to make educational decisions about students. A systematic approach to educational assessment is one in which a variety of assessment methods are used and in which all relevant assessment variables are considered.

### Assessment Methods

Information about a child's functioning in educational settings may be gathered through:

1. **Inspection of the child's record files.** Information reported in cumulative and other available records can be helpful in determining if there are factors that might account for problems the child is having, if there are trends in the growth of the problem, and if there are other relevant factors that need to be evaluated.

2. **Informal consultation with others who know the child.** Consultation with persons who know the child (e.g., peers, siblings, parents) may yield clues to further assessment.

3. **Structured interviews.** Interviews with the parent, the child, and/or teacher can yield information about the child's areas of difficulty and determine what resources have been used previously to collect information.

4. **Observation.** Through observation, assessment data is collected on the behavior of the student as it occurs naturally in the environment. Such data can focus on specific facets of behavior (e.g., child-teacher interaction), can indicate areas in need of further evaluation, and can confirm other information gleaned from records. Observations may be formal or informal and include anecdotal records, interaction analysis, checklists, and rating scales. Observation procedures can be both long-term and short-term processes.

5. **Norm-referenced tests.** Standardized tests can be used for both screening and diagnostic purposes to determine how one child compares with other children.

6. **Criterion-referenced tests.** This type of test measures a child's level of skill development in a particular area. It is especially useful for planning purposes because criterion statements can be used as goals for instruction.

## Relevant Assessment Variables

The decision-maker should gather educational information about the "whole" child. He/She should therefore consider all relevant factors which, because of their interactions with the child, may influence the child's functioning. These factors will vary with the child and the particular educational problems, but should include:

1. **Classroom environment.** The classroom environment refers to the physical arrangement of objects within the classroom (seating arrangements, materials, learning cen-

ters, etc.) which have the potential to impact learning, social interaction patterns, attention span, interests, achievement levels, etc.

2. **Curriculum.** A thorough knowledge of the curriculum presented to the child is necessary when making decisions about that child's educational functioning. The assessor should be aware of:

a. concept definition and sequence;
b. number of tasks and sequence;
c. evaluation component;
d. variety of instructional materials;
e. relevance of curriculum and materials to the child.

3. **Instructional style.** McIntyre (1978) has discussed the concept of "field-sensitive" and "field-independent" behaviors of classroom teachers. A "field-sensitive" teacher creates a warm, personal environment. He/she encourages the expression of students' feelings, relates concepts to the students' personal experiences, and sequences explanations from the general to the particular. He/she encourages and reinforces both group and individual efforts. The "field-independent" teacher is more formal and distant and prefers regularity, predictability, uniformity, and environmental control. While neither style is "good" or "bad," the nature of teacher interactions will affect student competence, behavior, and achievement. The effects may be limiting or demanding.

When making decisions about the functioning of students in educational settings, it is important to consider the three factors listed above. However, the primary focus during the assessment process should be on characteristics of the student(s).

Several models have been developed for systematic educational assessment. Underlying all models is the movement of the decision-maker from information which is general, broad, and only assumed to be accurate, to information which is specific, precise, and valid. The purpose of gaining valid information through the assessment procedure is twofold. First, the professional must be in a position to make valid intervention decisions about a pupil with special needs. Second, the assessment information obtained must provide the basis for an evaluation of the success of any interventions that are implemented.

In the following section, an example of a systematic assessment model is presented. Assessment methods and sources of information about the child are integrated into an assessment procedure.

## A Model for Systematic Assessment

### Screening

**Referral and verification.** The referral process is the first step in assuring the appropriate identification of children with special educational needs. Referrals are typically initiated by counselors, psychologists, classroom teachers or resource teachers. Information leading to the verification of educational problem(s) for which a child has been referred can come from examination of the child's cumulative records and direct observation of the child's behavior. Data on the child's personal, medical, and educational history from cumulative records can contribute to referral information and allow the decision-maker to discern the pattern of events leading up to the present situation. Observations can provide comprehensive, detailed, and specific information about the behavior(s) of the child and about the contexts or environments in which the observations are made.

When observing the child in interactions, the focus should be on the identification of habitually occurring events which reinforce the behavior being studied. Both social and nonsocial consequences which maintain the behavior must be identified, and an estimate of the child's responses to such stimuli should be made. It should also be possible to identify means of structuring the environment in the future. Once educational problems and target behaviors for the child have been identified, environmental contingencies which will reinforce the target behaviors can be determined.

One method of collecting the information outlined above is through a behavioral interview. While the behavioral interview is probably one of the least structured assessment strategies, it can nevertheless support the more objective data gathered through systematic observation. The following outline is suggested to ensure that all possible information is collected from the individual and that data relevant to the formulation of an achievement plan is gathered.

1. An initial analysis of the problem situation should be made, in which problematic behavioral excesses and deficits as well as nonproblematic behavioral assets are specified.
2. The problem situation should be clarified by identifying the individuals who object to the problem behavior and who may be affected by any behavior change made by the student. The clarification should also help to specify the conditions under which the behavior occurs.
3. A motivational analysis should be carried out, in which reinforcers (both positive and negatie) that may be maintaining the problem behavior or that may be useful in shaping more appropriate behaviors are specified.
4. A developmental analysis should be made, in which several questions are raised about the biological, sociological, and behavioral changes that may be pertinent to the problem behavior.
5. An analysis of the student's self-control should be completed, in which the limitations, conditions, and methods of self-control are defined.
6. Social relationships should be identified and the influence of significant others on the student specified.
7. The student's social, cultural, and physical environments should be examined, with considerations of cultural norms relating to the problem behavior, the similarity of norms in different settings, and various environmental restraints impinging on the student.

Schertzer and Linden (1979) also recommend that information be obtained from significant others in the child's life, e.g., parents and teachers. These second-party report measures help in obtaining information about the child's likes and dislikes (objects, events, places, activities, people, etc), and in identifying the perceptions of the teacher and parents as to the nature of the problem they see with the child. In addition, assessment instruments, such as the American Association on Mental Deficiency (AAMD) Adaptive Behavior Scale can assist in the gathering of second-party report information.

Appropriate screening procedures also require that information from previous assessments of the child's performance on standardized tests of general verbal and nonverbal ability (such as intelligence tests) and standardized educational achievement tests (such as reading, spelling, and math) be examined.

**Standardized achievement test data.** Achievement tests are devices that assess a student's skill development in academic content areas. Most tests are commercial instruments that sample the products of past formal and informal educational experiences and measure the extent to which an individual has profited from schooling and/or life experiences as compared to others of the same age or grade. The scores assigned to a pupil reflect both pupil ability and the instructional effectiveness of the educational program. When using information from achievement tests to decide if additional diagnostic assessment is necessary, the assessor must be aware (a) of the nature of behaviors sampled by the test, particularly the relationshp of specific curricula in use in the school district; (b) what specific items the student has passed and failed, and consistent failure patterns; and (c) that in the screening procedure, achievement tests determine, in a global way, only the child's current level of functioning.

Standardized achievement tests such as the Stanford Achievement Test and the Iowa Test of Basic Skills generally are used as screening devices. These tests in fact measure the extent to which a student has benefitted from past schooling compared to others of the same age or grade level. Knowing the nature of the subtests in these batteries allows some information to be obtained about the remediation needed by individual students and provides a general idea about where to start additional diagnostic assessment.

### Evaluation of Screening Data and Planning Interventions

In this step, the diagnostician evaluates the screening data and its reliability. It is important to be aware of the factors influencing reliability and validity of observation and test information.

Observational tools must be used with care and precision. Misinterpretation and misuse of observational techniques will distort estimates of individual behavior patterns which are at best only samples of behavior. Interpretation of information must include attention to potential problems of sampling error and of inadequate sampling. Observer bias and the possibility of the individual's awareness of being observed, in conjunction with the variability of behavior and the selectivity of observation, are threats to the validity and reliability of observational tools and techniques.

Information gained from interviews will vary among interviewers, and resulting inconsistencies mean that all interpretations must be viewed as tentative hypotheses to be verified or refuted with further evidence. For example, interview data from a teacher with limited skills, who is hoping to have a difficult child removed to a separate special education class, should be treated with caution.

**Factors affecting test validity.** Whether to use a test for a certain purpose and how to interpret the information yielded by the test are decisions which should be governed by the validity of the test. Validity is specific both to purpose and population. If the test is not valid, derived scores will be inaccurate. Despite efforts of test developers to ensure content, construct, and predictive validity, tests are only valid for certain purposes. Selection of standardized tests must be made with the purpose of the test in mind. These questions should be considered:

1. Was the test designed to be used for screening purposes?
2. Was the test designed to be used to place a student in a specific curriculum or in an educational program?
3. Was the test designed to be used to assist teachers and administrators in planning *individual* or *group* educational programs?
4. Was the test designed to be used for program evaluation?
5. Was the test designed to be used to assess individual progress?

**Factors affecting test reliability.** On another level, and related to the question of validity, are questions that must be asked to determine the accuracy of information yielded by the test:

1. Was the test technically adequate, that is, did it specify adequate information about test administration, standardization, reliability and validity?
2. Assuming that the test was technically adequate, was it used for the right purpose? For example, was a measure of receptive vocabulary, such as the Peabody Vocabulary Test, used as a measure of intelligence?
3. Assuming that the test was technically adequate, was it appropriate to the child? Did the student come from a cultural group comparable to those in the norming sample? Was the test age-appropriate? Did the child possess physical and mental characteristics comparable to those in the norming sample? Did the test require responses that a child with motor, visual, or hearing impairments would be unable to perform? Was the tester aware of any medication the child was taking? (Many drugs, such as Dilantin, Valium or Ritalin, may affect test results by influencing the nature and rate of responses.)
4. Is the assumption being made that norm-referenced scores such as grade and age equivalents, or composite scores such as IQ or SQ, are giving adequate representations of skills and deficiencies?
5. What information does the tester have about the child's background and current status and performance that will assist in accurate interpretation of test scores?
6. What relationship exists between the test and the curriculum in which the student is currently placed? Is the test measuring knowledge (and the chance to acquire it) or is it measuring variables that may not be in the student's repertoire of learned behavior?

If the professional cannot respond sufficiently to all of these questions, the test information must be interpreted with caution. Further diagnostic assessment is necessary to provide an accurate picture of the student's abilities and deficiencies so that an appropriate and relevant program can be developed.

### Intervention

The purpose of an educational diagnosis is to provide: (1) recommendations for instructional behavior management, and (2) recommendations for educational placement. In the following section, some methods for using assessment information for planning interventions will be presented.

---

*Task 2*

*Identify a child whom you have recently evaluated. How familiar were you with his/her classroom environment and with the curriculum into which he/she was placed? To what extent did you evaluate the child's teacher, his/her teaching style, and interactions with the child?*

---

## Norm-Referenced and Criterion-Referenced Assessment

Quality educational programming requires that testing instruments be used to ensure that each pupil enters a curriculum with the necessary prerequisite skills. A testing technology has been developed that is concerned with defining the relationship between an individual's skills and instructional sequences. Glaser (1963), in an article on the measurement of learning, defined two types of tests. One type, the norm-referenced test, evaluates the child's performance by comparison with that of the other children. The intelligence test is a classic norm-referenced test. The other type, the criterion-referenced test, evaluates the child in terms of some learning task at some absolute standard that is independent of other children's performances. The criterion-referenced test is designed to identify the relationship between a pupil and a specific instructional task. If a teacher wants to know what types or numbers of tasks a child has performed, then a criterion-referenced test would generally be called for.

Although a test is usually constructed as either norm-referenced or criterion-referenced, a person may use one test for both purposes. If we were to examine the items on a standardized achievement test for patterns of successes or failures on specific instructional tasks, we would be using it as a criterion-referenced test. If we used a test developed as criterion-referenced to rank the pupils in a class, we would be using it as a norm-referenced test because the child's score (rank) would describe him in terms of his relationship with others (the rest of the pupils in the class).

### Normative Testing and Instruction

The need to develop Individualized Education Programs (IEP's) for special education pupils requires reconsideration of the role of norm-referenced testing procedures. Normative tests are typically designed for such purposes as prediction, selection, and comparison. Special education has a long history of attempting to adapt these tests for individualized instruction. Except for a few standardized diagnostic tests which have both norm-referenced and criterion-referenced properties, the majority of norm-referenced tests provide limited information for guiding the teaching of specific skills. A major problem encountered in the use of norm-referenced tests is that most are not designed to relate directly to a specific curriculum sequence. The discrepancies between norm-referenced and criterion-referenced test information become especially important when using such information for individual program planning.

**Table 1**

**Two Types of Testing Compared**

| **Norm-Referenced** | **Criterion-Referenced** |
|---|---|
| • *Evaluates individual performance in comparison to a group of persons.* | • *Evaluates individual performance in relation to a fixed standard.* |
| • *Is used to evaluate a student as "below grade level," "at grade level."* | • *Not concerned with grade level descriptions.* |
| • *Fails to indicate where students have mastered the spectrum of instructional objectives.* | • *Identifies the individuals who have mastered the spectrum of instructional objectives.* |
| • *Generally poor aides in planning specific instructional programs.* | • *Geared to provide information to be used in planning instruction.* |
| • *Is often vague in relation to instructional content.* | • *Is content specific* |
| • *Does not operationally define mastery and/or success.* | • *Operationally defines mastery and/or success.* |
| • *Applies poorly to the individualization of instruction.* | • *Applies directly to the individualization of instruction.* |
| • *Is not concerned with task analysis.* | • *Depends upon task analysis.* |
| • *Test results interpreted in references to a person's position in relation to the scores of others.* | • *Test results interpreted in reference to a person's position in relation to the curriculum.* |

*—adapted from Housden, J. L., & LeGear, L. An emerging model: Criterion-referenced evaluation.* **Thrust,** *April 1975.*

In the example in Table 2, data were collected from twenty-four pupils referred for special education service as spelling failures. Two types of test data were collected: (a) norm-referenced data from a standardized test, and (b) data from a criterion-referenced test that identified the placement of a child with regard to the specific spelling curriculum in use in the school. Generally, a criterion-referenced test is not used to identify grade levels because this usually entails moving from specific practical information to more general and less prescriptive information. For purposes of comparison in this case, however, a grade level score was assigned to the criterion-referenced data. This grade level score indicated the level in the curriculum set for that grade at which instruction should begin. An observation of the scores in Table 2 reveals that the difference between the overall mean grade level scores on the two tests was not great (0.14). However, when the pupils are considered as individuals, relationships between normative scores (standardized achievement test) and criterion tests (curriculum placement test) are so great that the rather common practice of using standardized achievement test data for individual program planning must be questioned. In analyzing the content of the two tests, it was found that the criterion-referenced test content was consistent with the specific spelling program in use in the school, which placed a strong emphasis on controlled sequences built around the phonetic structures of words. In this instructional program, irregular words were emphasized much later than they are in some other types of spelling programs that use different theoretical rationales such as "frequency of use." The content of the standardized achievement test stressed irregular words much earlier than did the criterion-referenced test; it was, therefore, of little value for making decisions related to the phonetically oriented curriculum being used for the twenty-four children listed in Table 2. Thus, it is important to choose tests that relate to the curriculum being used in order to obtain accurate data in developing the child's individual program.

**Table 2**
**SPELLING FAILURE REFERRALS**
**Norm-Referenced Score—Curriculum Placement**

| | Ach. Test Score* (Norm-referenced) | Curr. Placement* (Crit.-referenced) | Differences in Placement* | |
|---|---|---|---|---|
| | 1.1 | 2.0 | - .9 | |
| Grade | 2.0 | 2.2 | - .2 | |
| 3 | 2.0 | 2.2 | - .2 | X = -.45** |
| | 1.7 | 2.0 | - .3 | |
| N=6 | 1.7 | 2.0 | - .3 | |
| | 1.4 | 2.2 | - .8 | |
| | 1.8 | 2.2 | - .4 | |
| Grade | 1.8 | 2.0 | - .2 | |
| 4 | 2.1 | 2.2 | - .1 | X = -.23** |
| | 1.5 | 2.2 | - .7 | |
| N=6 | 1.5 | 2.0 | - .7 | |
| | 2.9 | 2.2 | + .7 | |
| | 3.2 | 2.8 | + .4 | |
| Grade | 3.2 | 2.2 | +1.0 | |
| 5 | 3.2 | 2.2 | +1.0 | X = .55** |
| | 3.4 | 2.2 | +1.2 | |
| N=6 | 3.2 | 2.8 | + .4 | |
| | 1.9 | 2.6 | - .7 | |
| | 4.5 | 3.4 | +1.1 | |
| Grade | 2.9 | 2.8 | + .1 | |
| 6 | 2.9 | 3.4 | - .5 | X = .68** |
| | 4.0 | 2.8 | +1.2 | |
| N=6 | 4.8 | 2.6 | +2.2 | |
| | 2.9 | 2.8 | + .1 | |

Total number of subjects = 24
Differences between means = .14
Correlation between normative and criterion-referenced test scores = .65
* All scores expressed in grade levels.
** Positive score indicates normative test higher.

## The Process of Diagnosis

The process of diagnosis in an instructional setting has been likened by some to the medical process from which the term has its roots (i.e., a doctor examines the patient's symptoms and the pattern of these symptoms leads him to the *diagnosis* that a certain disease or dysfunction is present).

This analogy is spurious, however, for it implies that the learning diagnosis identifies or defines a knowledge "disease" or a disorder rather than a specific need or a cluster of needs, which is what actually takes place. Diagnosis in education, then, is actually a *needs assessment process*.

The instructional manager uses his professional skills and the instruments and techniques available to him to examine the present state of the learner in relation to pre-established learning outcomes or objectives, usually arrayed in a continuum or hierarchy of progression. (Hickey & Hoffman, 1973, p. 36).

### Diagnostic Devices

There are a number of standardized diagnostic tests available that have both norm-referenced and criterion-referenced properties. Many standardized diagnostic tests can tell the teacher whether a child is above or below average for his chronological age in the area tested, and in sub-areas that make up the general area. These tests have value for program evaluation, particularly if accountability data are required by an outside agency. Many outside evaluators are reluctant to accept data from locally developed criterion-referenced measures. In such cases, pre- and posttesting with standardized diagnostic tests may provide a more acceptable approach. In most special education situations, standardized tests will have to be used in combination with criterion-referenced tests that are more specific to the local curriculum and associated materials. A standardized diagnostic math test might indicate that a child is weak in addition combination, but a more specific instrument will be needed to determine which addition combinations need to be taught to the child. Some diagnostic tests which are commonly used are those that assist in pinpointing specific strengths and weaknesses in reading and mathematics.

According to Salvia and Ysseldyke (1978) diagnostic tests of reading typically lack reliability and validity. Therefore, diagnostic assessment of reading skills should consist of repeated measurements to confirm or disallow initial patterns of error.

The skills assessed by diagnostic reading tests include those of oral reading, comprehension, word attack, word recognition, and rate of reading. Tests commonly used are the Gray Oral Reading Test, Gates-McKillop Reading Diagnostic Tests, Durrell Analysis of Reading Difficulty, Stanford Diagnostic Reading Test, and the Woodcock Reading Mastery Test. Another diagnostic test, the Fountain Valley Support System in Reading, provides criterion-referenced information, whereas other tests may require the decision-maker to determine the pattern of errors and correct responses. This system includes pupil profiles which provide a record of individual pupil achievement and enable monitoring of pupil progress.

Diagnostic assessment in mathematics leads to specific information about a student's performance in content areas (e.g., numeration, fractions, algebra) operations, and applications (e.g., measurement, time, problem-solving, money). Three most commonly used diagnostic mathematics tests are the Key Math Diagnostic Arithmetic Test, the Stanford Diagnostic Mathematics Test and Diagnosis: An Instructional Aid in Mathematics.

### The Relationship of Domain-Referenced Testing to Criterion-Referenced Tests

The quality of a criterion-referenced test depends on the degree to which a skill area is clearly identified and represented within the test. Some criterion-referenced tests are poorly constructed. In recognition of this fact, Hively (1974) coined the term "domain-referenced." A domain-referenced test is one in which the emphasis is placed on precisely identifying skill areas (domains). The test items of a domain-referenced test are selected so that one can be certain that a child who meets criteria on the items which represent a domain would be able to master all other possible items of the domain when they were encountered. For example, if a series of test items was prepared to test a child's mastery of a domain, such as long division of decimal fractions using single digit divisors, the test items would have to include examples with zero in the dividend. Without such examples involving zero, there would be little guarantee that performance on the test items would generalize to other long division examples a child might encounter.

In describing the term "domain-referenced," Donlon (1975) noted:

> The label "criterion-referenced" has what Hively calls "surplus associations." Further, Hively has recognized that our educational objectives are seldom detailed: We say we want the child to "know the alphabet," not to know the letter "a." We aim, then, at classes with related behavior; the model is not one of appraising the "ability to jump from standing"—a domain of jumps, forward and back, sideways, landing on one foot, and so forth . . . Hively and his associates have thus improved upon "Classical" criterion-referencing. They have stressed the complexity of domains by pointing out subdomains. . . . (p. 39)

Domain-referenced testing is one approach to criterion-referenced testing in which considerable care is taken with test item selection to ensure that we can generalize from performance on the test items to the specific curriculum area (domains) that the test items are selected to represent.

*Criterion-referenced testing can reach its full potential only when it is so integrated into the day-by-day functioning of the classroom that it cannot be separated out as a "testing" activity.* Indeed, its contribution to the direction and programming of instructional activities should be such that the teacher sees it as indispensable for facilitating effective instruction.

The diagnostic class profile is a basic classroom management component of a criterion-referenced test. Unlike many "clinical" tests, a school oriented criterion-referenced test has to take into account the management problems faced by the teacher who is responsible for the total class. The basic parts of a diagnostic class profile are (a) a listing of the pupils on one axis of a matrix, and (b) a listing of the specific skill being tested on the other axis of the matrix.

An example of a diagnostic class profile used with a criterion-referenced punctuation test (Hofmeister, 1972) is listed in Figure 1. Along the top of the profile are 20 skills in a suggested order of instruction from left to right. A listing of the pupils whose skills are being evaluated is on the left hand side of the profile. The top right-hand corner of the profile lists the coding procedure used to record test data. Most class profiles can be adapted to monitor progress and to provide pre- and posttest data.

A well structured criterion-referenced test should generate direct individual prescriptions for each child. In the case of the punctuation test (Figure 1), the individual prescriptions are obtained by reading from left to right after the child's name. Small group prescriptions can also be obtained by working vertically down the class profile.

## FIGURE 1. DIAGNOSTIC CLASS PROFILE PUNCUATION TEST

↘ Passed

– Failed

↘ Retested Passed

| | Initial capital letter | Terminal sentence period | Captialize "I" | Question Mark | Capitalize proper nouns | Capitalize days, holidays | Commas in a series | Capitalize titles | Initials and abbreviations | Quotation marks | Commas to separate clauses | Quotation and commas | First word of quotation | Streets as a name | Parenthesis to enclose | Nonrestrictive phrases | Contractions | Adverbial clause | Possessive apostrophes | Colons | Total Errors Possible | Student's Score | NOTES |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Punctuation Rules | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | | | |
| Possible Errors Level 1 | 4 | 4 | 2 | 3 | 6 | 2 | 6 | | | | | | | | | | | | | | | | |
| Possible Errors Level II | 5 | 4 | 5 | 4 | 15 | 5 | 12 | 4 | 4 | 12 | 5 | 6 | 3 | 3 | 8 | 12 | 4 | 2 | 5 | 5 | | | |
| Pupils: 1) | | | | | | | | | | | | | | | | | | | | | | | |
| 2) | | | | | | | | | | | | | | | | | | | | | | | |
| 3) | | | | | | | | | | | | | | | | | | | | | | | |
| 4) | | | | | | | | | | | | | | | | | | | | | | | |
| 5) | | | | | | | | | | | | | | | | | | | | | | | |
| 6) | | | | | | | | | | | | | | | | | | | | | | | |
| 7) | | | | | | | | | | | | | | | | | | | | | | | |
| 8) | | | | | | | | | | | | | | | | | | | | | | | |
| 9) | | | | | | | | | | | | | | | | | | | | | | | |
| 10) | | | | | | | | | | | | | | | | | | | | | | | |
| Total Possible | | | | | | | | | | | | | | | | | | | | | | | |
| TOTAL | | | | | | | | | | | | | | | | | | | | | | | |

Class ________________ School ________________ Teacher ________________ Date ________________

### Adapting Test Content

The content of many tests is often accepted as unalterable because of association with normative testing in which alteration of either the test content or the administration procedures results in invalidating any normative information calculated from such administrations. In some criterion-referenced test situations, alterations can be made without invalidating the test. In the math combinations test (Figure 2), a number of the examples require negative numbers as answers. If a teacher feels that these negative number problems are inappropriate for her needs, she can delete such problems to make the test more responsive to her needs.

### Random and Linear Access

In the math combinations test (Figure 2), there is very little inherent sequence among the subtraction examples. Because of that fact and the relatively low conceptual level of the task, it makes little difference where one begins teaching. This test is an example of a random access test.

There are other test areas where sequence is important, as with the sub-skills of the long division algorithm, or a spelling list based on word difficulty. Tests in which sequence is important are termed "linear access tests." The designation of a test as a linear or random access test has a number of instructional and testing implications, such as adaptability to the special educational or physical needs of the child. In a random access testing situation, flash cards can often be used both for testing and teaching. The use of a set of flash cards for each child is one of the quickest and most practical ways to identify and group the items a child knows and does not know. In a linear testing situation, it is not always necessary to test all the skills in the test area. For example, in spelling it mig[illegible] sufficient to test until the child starts failing and ther[illegible] teaching. Further testing after the child has started failing would certainly be aversive to the child, and very likely a waste of instructional time.

### Cross-Referencing

Prescriptions based on the results of criterion-referenced tests, where possible, should cross-reference from skills to instructional materials. Skills might be cross-referenced as shown in the following example. In the example, the page numbers refer to the text being used in the child's classroom.

| Skills | Instructional Examples Page Number | Test Examples Page Number |
|---|---|---|
| 1. parts of a mixed number | 172-173 | 174, 192, 456 |
| 2. mixed number products | 176 | 184, 192, 456 |

### Commercial Criterion-Referenced Testing Systems

A number of companies publish comprehensive testing systems which provide data on a child's performance on several hundred skills. Some of the systems require computer scoring services. Some of these testing systems can be cross-referenced with most of the major basal reading programs and math texts in use in the schools. Examples of some of the available criterion-referenced testing systems are: The Prescriptive Mathematics Inventory, published by the California Test Bureau, A Division of McGraw-Hill Book Co.; The Fountain Valley Teacher Support System, published by Zweig Associates; The Wisconsin Design Program, published by NCS Interpretive Scoring System; and The Beginning Assessment Test for Reading, published by Lippincott.

The following description from the preface to the Teacher's Handbook for *Diagnosis: An Instructional Aide - Mathematics, Levels A and B,* (developed and published by Science Research Associates, 1980) is representative of the approach taken in some of the well-developed commercially available criterion-referenced testing systems:

> Each diagnosis lab provides sets of diagnostic probes—exercises designed to help the teacher identify the specific weaknesses of a student. The exercises are based on the comprehensive lists of learning objectives on the back of the probe. By freeing the teacher from the time-consuming task of developing detailed diagnostic tests, a diagnosis lab enables him to focus his efforts on the instruction or remediation needed by each student to correct identified weaknesses.
>
> Also provided is a Prescription Guide with which teachers and students can quickly determine where—in a large number of texts and supplementary learning kits—materials pertinent to the achievement of each objective are presented. Multiple references are given for every learning objective. Thus, if the student has had difficulty with a given text presentation or other instructional material, he/she can be directed to another. This approach allows for a multitext and multimedia approach, accommodating a wide variety of learning and teaching styles.
>
> While some educators still champion "tacit learning" and stress the impossibility of reducing all aspects of the educational process to behavioral objectives and demonstrable outcomes, it is generally accepted that mastery of fundamental skills is, perhaps, best taught in this fashion. This is the basic premise of the series of diagnostic labs; to provide a convenient method of helping students attain all the fundamental skills and concepts upon which much of their future education will rest. (p. 1)

### Setting a Criterion

One of the major purposes of a criterion-referenced test is to help determine which skills to incorporate in remedial instruction. The manual accompanying a criterion-referenced test should specify skills to be tested in a sequence to which skills outlined in a curriculum can be fitted if necessary. As yet, not enough research has been done to indicate precisely what degree of success should be used to indicate when a skill is mastered. The following are guidelines for setting a criterion:

1. **A criterion should be based on the subject matter,** and one level of obtainment (e.g., 95 percent correct) should not be expected to apply to all subject matter areas.

2. **The range used should be 80 percent to 100 percent.** If a skill is important in terms of personal safety (for instance, driver's training) or as a prerequisite to an important skill, a range of 90 percent to 100 percent would be appropriate.

3. **The criterion should be related to the nature of the test.** If the test item was designated as follows:

Question (1) $\begin{array}{r} 34 \\ -19 \\ \hline \end{array}$

then a criterion of 90 percent on the total test might be appropriate. If the test questions were of the form:

Question (1) $\begin{array}{r} 34 \\ \underline{-19} \end{array}$ Check one: (a) 13

(b) 14

(c) 15

(d) 16

then a criterion of 100 percent might be more appropriate because of the chance factor and prompting associated with the multiple choice format.

4. **A criterion may take into consideration the performance of others.** A level may be established by referring to the relative position the student holds in a particular group. This adds a norm-referenced element to the test.

5. **A level may be set by judging minimal competence.** Experts decide what score a minimally competent person should obtain.

**Subtraction**

## DIAGNOSTIC ARITHMETIC

**Date:** **Grade:** **Time:** **Name:**

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 6<br>-4 | 13<br>-6 | 6<br>-2 | 0<br>-6 | 5<br>-1 | 1<br>-0 | 15<br>-8 | 7<br>-3 | 5<br>-5 | 14<br>-7 | 13<br>-9 |
| 15<br>-6 | 5<br>-2 | 3<br>-2 | 6<br>-1 | 4<br>-0 | 11<br>-8 | 3<br>-3 | 11<br>-5 | 12<br>-7 | 10<br>-9 | 13<br>-4 |
| 4<br>-2 | 6<br>-8 | 4<br>-1 | 9<br>-0 | 13<br>-8 | 9<br>-3 | 14<br>-5 | 7<br>-7 | 14<br>-9 | 8<br>-4 | 14<br>-6 |
| 0<br>-2 | 8<br>-1 | 2<br>-0 | 12<br>-8 | 8<br>-3 | 12<br>-5 | 16<br>-7 | 9<br>-9 | 7<br>-4 | 9<br>-6 | 10<br>-2 |
| 2<br>-1 | 8<br>-0 | 16<br>-8 | 6<br>-3 | 13<br>-5 | 8<br>-7 | 11<br>-9 | 5<br>-4 | 8<br>-6 | 9<br>-2 | 2<br>-11 |
| 6<br>-0 | 17<br>-8 | 8<br>-3 | 7<br>-5 | 11<br>-7 | 16<br>-9 | 10<br>-4 | 11<br>-6 | 8<br>-2 | 3<br>-9 | 10<br>-1 |
| 8<br>-8 | 5<br>-3 | 8<br>-5 | 15<br>-7 | 12<br>-9 | 11<br>-4 | 7<br>-6 | 3<br>-2 | 0<br>-4 | 7<br>-1 | 5<br>-0 |
| 10<br>-3 | 6<br>-5 | 10<br>-7 | 18<br>-9 | 12<br>-4 | 6<br>-6 | 11<br>-2 | 2<br>-6 | 9<br>-1 | 0<br>-0 | 14<br>-8 |
| 9<br>-5 | 13<br>-7 | 17<br>-9 | 9<br>-4 | 12<br>-6 | 2<br>-2 | 0<br>-6 | 3<br>-1 | 7<br>-0 | 10<br>-8 | 4<br>-3 |
| 9<br>-7 | 15<br>-9 | 4<br>-4 | 10<br>-6 | 7<br>-2 | 8<br>-11 | 1<br>-1 | 3<br>-0 | 9<br>-8 | 12<br>-3 | 10<br>-5 |

Figure 2. Math combinations test in subtraction with diagonal sequences: "Diagnostic Arithmetic."

---

Task 3

*a. Identify a norm-referenced test which you commonly use. Evaluate the test against the statements about norm-referenced tests in Table 1. Have you used the test to make placement decisions? If so, do you still consider it to be an appropriate instrument for this purpose?*

*b. To what extent have you used (or recommended) criterion-referenced tests for verification of your diagnosis? If not, on what information have you based your recommendations for remedial instruction? How successful have these recommendations been?*

---

### Use of Profiles in Planning Educational Programs

Teachers and psychologists are often interested in the relative levels of mastery that a student demonstrates in various skill areas, or in the differences between performance in domains within one test. Although the construction of profiles is common practice, their analysis and interpretation are complex procedures. The professional must be cautious in using such information for purposes of classification and program planning.

The assumption is generally made that abilities and skills in educational settings develop at a consistent rate and that the relative "flatness" of a child's profile (all scores falling within 1.6 standard deviations of the mean) allows the diagnostician to make decisions about the "normalcy" of performance.

Salvia and Ysseldyke (1978) note that "educators expect children with above-average intelligence to perform better than average in their academic work. If their achievement is not also above average, it is often a source of concern" (p. 408). It is also true that children who achieve relatively low scores on intelligence tests are expected to perform below average. Too often, poor performance on a test of ability leads to expectations of poor performance in other skill areas.

Professionals have used assumptions about the correlation between ability and achievement to support classification of students in different areas of exceptionality. Salvia and Ysseldyke (1978) point out that:

> Flat profiles of individuals whose functioning measures significantly below average in both intelligence and adaptive behavior are used to confirm diagnoses of mental retardation. . . . A child who has a significant discrepancy between measured intelligence and both measured achievement *and* perceptual or language functioning, or both, may have a learning disability. (p. 409)

It is also assumed that program planning may be facilitated through use of profiles of achievement tests. Poor performance in one or more academic areas may indicate a need for additional instruction. Conversely, high achievement in an academic area may indicate special interests or skills which can be capitalized on in instruction.

If profile analysis is to be of use in developing instruction, attention must be given to the reliability of score differences and to the differences in norm samples. Salvia and Ysseldyke (1978) provide an example of profile analysis which is reproduced in Appendix C.

## Evaluation of Intervention Effects

The success of an intervention program on behalf of a pupil with special needs should be evaluated from two aspects: (1) the extent to which the treatment has resulted in significant change and/or progress, and (2) the extent to which the treatment program is valid.

### Evaluation of Intervention Outcomes

Public Law 94-142 requires both assessment of students who are in need of educational interventions and the evaluation of the effectiveness of these interventions in achieving objectives for the students. The current emphasis on individual educational plans (IEP's) has resulted in a need for the professional to become involved with assessment of outcomes of interventions. As a member of a team which develops an IEP, the school psychologist must be prepared to take part in the review of treatment effectiveness.

In his discussion of the assessment of behavioral change, Phye (1979), points out that the educational professional is concerned both with the nature of the change and with the evaluation of the impact of the intervention.

**Nature of the Change.** Here, the types of questions to be asked deal with the reliability of the data showing a change in behavior and the direction of the change. Evaluation in this area is typically concerned with the improvement and/or modification of the intervention program. A standard method of assessing the nature of a change due to intervention is through a simple pretest-posttest procedure. In addition, the assessment of instructional objectives must be dealt with on a continuing basis at all levels of an instructional task in order to monitor instructional methods and materials.

**Evaluation of Intervention Impact.** To evaluate the effectiveness of the intervention program the following question should be asked: "To what extent was the observed change a result of the intervention treatment?" Here the program to be evaluated may be an IEP or a program or curriculum through which treatment is prescribed.

Given the ethical and practical restrictions that are placed on the professional in the school setting, evaluation of program effectiveness via "true" experimental design is not generally possible. Phye (1979) therefore suggests that some form of time-series design would be valuable in ascertaining intervention impact.

**Treatment Validity.** Closely related to the issue of evaluating program effectiveness is that of treatment validity. It is common practice in special education to use certain tests as a basis for prescribing educational treatment. When tests are used in this way, it is necessary to assess their treatment validity, i.e., to assess the degree to which they enable educators to accurately describe a treatment effect. In this section a procedure is described for assessing the treatment validity of a test.

There are essentially three areas of possible threat to treatment validity. These areas were described by Gallery and Hofmeister (1978) as: (a) the lack of a relationship between test and treatment; (b) the low quality of treatment, and (c) the lack of relationship between treatment and curriculum.

**Relationship Between Test and Treatment.** Can an educational treatment be prescribed, given the test results? If the answer is no, decisions about treatment are unreliable. The treatment decisions may result in the student's failure to achieve, and such failure would be difficult to assess since it could be the result of inaccurate measurement, ineffective treatment, or both. Related to this concern is the teacher's ability to determine a starting point for treatment. If the test results do not give some indication of where the treatment should begin, the treatment may be inefficient or ineffective.

**Quality of Treatment.** Was the treatment field-tested? There must be empirical evidence for an assumption that the treatment is logically and/or conceptually sound. Empirical scrutiny of the treatment should yield affirmative answers to the following questions:

1. Is there evidence that the skills contained in the treatment were mastered by students on whom the materials were field tested?
2. Are the responses required in the test similar to those required in the treatment?
3. Are the field-test students similar to the pupil(s) being tested?

**Relationship Between Treatment and Curriculum.** To what extent are the treatment objectives contained in the curriculum? The issue here is one of efficient use of time. A discrepancy between treatment objectives and the curriculum content could mean that the treatment has little relevance to the student's needs.

Validity involves all interpretations of the test data. Validity does not necessarily refer to the nature of the test(s) being used. That is, one cannot automatically assume treatment validity for a criterion-referenced test and not for a norm-referenced test. The validity of a test and the extent to which test results can assist a professional concerned with prescribing treatment depend on the degree to which the criteria outlined are met.

## Interpretation and Communication of Assessment Information

The following is a summary of protections relating to communication and assessment under P.L. 94-142:

1. Parental permission must be obtained before any initial diagnosis or evaluation can be started relative to the development of an IEP or for placement within a service pattern that is beyond that found within a regular education classroom.
2. Evaluation must be conducted by a multidisciplinary team of professionals designated by the local educational authority.
3. There must be evidence of due process or procedural safeguards in which the child can contest being placed in or being kept from attending a specially designed educational program.

The role and responsibility of the school psychologist is to gather reliable and valid information about a student's level of educational functioning and his/her educational skills and deficits. It is also the psychologist's responsibility to interpret and communicate that information in an understandable and educationally relevant manner. Shertzer and Linden (1979) suggest that persons who need to receive, understand, and use assessment information are school professionals (teachers, counselors, therapists, social workers, school administrators, and other members of the local education agency), the student's parents (or representative) and the student (if appropriate). The individual needs of each member of this audience are different but not mutually exclusive.

### Administrators

Administrators need information from norm-referenced tests to make decisions for (a) placement of students in educational programs; (b) evaluation of educational progress; (c) appraisal of the effectiveness of specific curricula and programs; and (d) identification of students with specific disabilities for funding purposes.

### Teachers

Teachers need assessment information that will assist them in knowing what to do instructionally. It is not sufficient to know that the student deviates from the norm. It is necessary to have specific information about skill development, strengths, and weaknesses to make decisions about effective intervention strategies.

Teachers and diagnosticians need to examine the relationship between the behaviors assessed by tests and the behaviors about which they need information. This information can be derived from interpretation of norm-referenced data in conjunction with the results of criterion-referenced tests.

**Parents**

Parents have the right to know whatever the school knows about the abilities, the performance, and the problems of their children. The school psychologist has the obligation to communicate information in an understandable and usable way. Lien (1971) lists several principles to follow in interpreting data to parents and students:

> 1. Make sure that both the professional and the person to whom test results are interpreted have a clear immediate goal in mind that serves as a reason for the interpretation.
>
> 2. Avoid the use of specific scores wherever possible. To give the parent a child's IQ score, for example, is not an appropriate practice. It is likely to be interpreted as a fixed characteristic and as a final conclusion about the child. A description of percentile scores is appropriate, but some explanation should be given as to what the score means and to whom the student is being compared.
>
> 3. Never discuss the implications of scores in terms of absolute answers. Combine the report with information about the test or battery and the relationship of the student's performance to that of others who have taken the tests.
>
> 4. Remember that understanding and acceptance are not synonymous. (p. 259)

**Students**

If the student is considered capable of understanding and making use of the assessment information to improve his/her performance, he/she has the right to be informed of test results and their implications. Once again, the information must be communicated in an understandable way, with attention to both content and language.

---

### *Task 4*

*a. To what extent do you typically conduct a systematic evaluation of the effect of instructional outcomes for those students with whom you have interacted? How do you assist classroom teachers in determining the significance of changes that have taken place?*

*b. To what extent do the tests you commonly use interface with valid treatment prescriptions? How relevant are tests to your school's (or district's) curricula?*

*c. Evaluate the nature of your interactions with administrators, teachers, parents, and students against the recommendations outlined in this section.*

---

## Conclusion

This overview of educational assessment and its purposes has been designed to assist the progress of the diagnostician through the complex maze of educational decision-making.

Educational assessment is a multidimensional procedure by which the infinite variety of a child's behaviors is observed and evaluated. Legal and social mandates combined with research findings delineate the responsibilities of school professionals in assessment. These requirements have resulted in a crucial need for the educational decision-maker to understand the nature and limitations of assessment technology.

The goal of educational assessment is the generation of educationally relevant decisions designed to make positive differences in the lives of children. The school professional must be prepared to be accountable for such decisions. One measure of such accountability will be the valid and reliable relationship between assessment, the demonstrated needs of the student, and the success of instructional strategies designed to provide for those needs. To assess the pupil and ignore the instructional environment is to imply that the pupil is totally responsible for the level of education.

## POSTTEST

1. According to P.L. 94-142, the primary purpose of assessment is to
    a. develop appropriate educational programs for handicapped children.
    b. make educational decisions about the placement of handicapped children in curricula and programs.
    c. evaluate the effectiveness of the child's individual education program in meeting identified goals and objectives.
    d. all of the above.
2. Because school professionals have different training and unique skills, the responsibility for educational assessment should be
    a. allocated to the person who spends the most time with the child in educational settings.
    b. the sole responsibility of the school psychologist as a function of his/her training.
    c. an integrative process involving every person who has contact with the child in the educational setting.
    d. the co-equal responsibility of those who are in a position to contribute educationally relevant information about the child.
3. Continual evaluation of the child's progress is necessary for the following reason(s):
    a. to make sure the assessment instruments used reliably identified the general nature of the child's problems.
    b. to ensure that the child is not causing further problems for the teacher.
    c. to ensure that the student is in the right curriculum and is learning at his/her maximum rate.
    d. to ensure that the student's placement will be final and that there will be no threat of his/her personal adjustment because of movement to a new position
4. In using the information from observation and interviews, the professional
    a. can view the information as accurate because of the reliability and validity of the instruments used.
    b. knows that he/she established rapport with the individual and, therefore, has obtained accurate information.
    c. knows that the information he/she obtained is more accurate than information from the child's teacher because the latter is not trained to observe.
    d. knows that interview-observation techniques are too subjective to yield precise and reliable information, but can be used to supplement and verify data gathered on the child.

5. A criterion-referenced test places a child in terms of
   a. a sequence of tasks.
   b. a set of percentiles.
   c. other children.
   d. a set of norms.

6. Norm-referenced tests are often
   a. sensitive to a specific curriculum.
   b. good diagnostic instruments.
   c. insensitive to a specific curriculum.
   d. useful for directing specific remedial programs.

7. A domain-referenced test is one form of a
   a. standardized diagnostic test.
   b. norm-referenced test.
   c. standardized achievement test.
   d. criterion-referenced test.

8. The intelligence test is a
   a. norm-referenced test.
   b. criterion-referenced test.
   c. domain-referenced test.
   d. achievement test.

9. If a criterion-referenced test was used to rank the pupils in a class, the test would be used as
   a. a criterion-referenced test.
   b. an achievement test.
   c. a norm-referenced test.
   d. a domain-referenced test.

10. There is a strong movement to de-emphasize testing for purposes of
    a. identifying and classifying.
    b. remediation.
    c. directing instruction.
    d. diagnosis and remediation.

11. The most useful special education testing procedures are those procedures which
    a. give a percentile ranking.
    b. facilitate remediation.
    c. facilitate giving end-of-term grades.
    d. place a child in relation to his peers.

12. When a child meets criteria on a domain-referenced test, he should
    a. have mastered the test items only.
    b. be able to master all examples of the domain when encountered.
    c. have mastered at least half of the possible examples of that domain.
    d. none of the above.

13. When selecting a test, the professional should first
    a. decide on the scoring procedure.
    b. make sure that there are at least two forms of the test.
    c. clarify the purpose of the test.
    d. make sure the students can understand the directions.

14. For criterion-referenced testing programs to be effective,
    a. they should be administered by an impartial outsider.
    b. they should be administered at the end of the year.
    c. they should be integrated into the daily teaching procedures.
    d. they should be administered at the beginning of the year.

15. A matrix with one axis listing the skills and the other listing the pupils is a
    a. table of norms.
    b. diagnostic class profile.
    c. individual prescription.
    d. standardized test.

16. A criterion-referenced test can be made more practical by
    a. adding norms.
    b. expressing results as a percentage.
    c. cross-referencing the skills.
    d. expressing results as ranks.

17. A diagnostic class profile can be used for
    a. recording pretest data only.
    b. recording pre- and posttest data and monitoring student progress.
    c. recording posttest data only.
    d. recording the data on which each class is given.

18. A well-constructed criterion-referenced test should
    a. generate prescriptions for direct individual instruction.
    b. generate percentile rankings.
    c. provide information regarding the appropriate norm-referenced test to give.
    d. none of the above.

19. If a teacher feels that parts of a criterion-referenced test are inappropriate, he/she should
    a. not use that particular test.
    b. adapt the test by deleting the parts he/she feels are inappropriate.
    c. make up his/her on test.
    d. delete those parts he/she feels are inappropriate, but realize that the test is no longer valid.

20. Test developers claim that Test A is more reliable than Test B. To determine the accuracy of their claim, what information would you need about Test A?

21. The authors of the WRAT state that no attempt was made to obtain a representative national sample of students for the standardization of the test. Each level of the test was standardized on at least 150 males and 150 females at each of nineteen age levels, producing a total standardization population of 5,868 persons at Level I and 5,933 persons at Level II. Norms were not stratified on the basis of race, ethnic group membership, socio-economic level or geographic region. Schools in only seven states were included in the standardization sample: No handicapped children were included (Salvia and Ysseldyke, 1978, p. 160).
    Identify five characteristics students may possess that require caution in interpreting test results and that may demand adaptation in test administration.

22. What are the two questions to ask when evaluating the nature of behavioral or academic change in a student during and after intervention?

23. When assessing treatment validity, which three areas will be of concern as possible threats to validity?

24. How do the needs of administrators, in regard to assessment information, differ from the needs of teachers? What are the practical implications of these differing needs?

## SIMULATION EXERCISES

1. It has been noted in the research literature that the lack of mastery of the sounds of letters is a major skill deficit in many poor readers. You will be presented with a paragraph which a student has read with markedly poor skill. Specific problems are identified in the test by a code which appears beneath the paragraph. Look for a consistent pattern and frequency of errors. Design a criterion-referenced test which will cover all the problems identified. The test will be individually and orally administered as a pre- and posttest to monitor instruction in the pronunciation deficits.

   Also prepare a diagnostic class profile to be used in association with the test. The class profile should include the following components:
   a. a space for listing at least 10 pupils;
   b. a listing of the sound being evaluated; and
   c. a clear and simple coding procedure for entering the test results. The coding system should allow for updating the profile as skills are mastered.
   All the administrative procedures should be described in sufficient detail so that another teacher might use the instrument without additional information and/or training.

### Reading Paragraph
### VAN'S CAVE

Van is an old, *old* man. He lives in a cave near a lake. Van has a bed and a stove in the cave. Van made the bed of old lumber. He made the stove of old bricks. He sleeps in the bed. Spot likes to sleep near the stove. His bed is just an old coat.

Spot likes to live in Van's cave. It is home for the old man and the dog. Once Van and Spot went to hunt ducks. Van saw a flock of ducks land on the lake. He fired and fired at the ducks. "I hit five ducks," said Van. "Jump into the water, Spot, and get the ducks for me." Spot jumped into the waves and swam to get the ducks.

**Scoring Code**

1. (letter or word sounded incorrectly)
2. (word skipped)
3. c (word or letter which student self-corrects)
4. (insertion)
5. sk (rows skipped)

2. You will be presented with a case report of a student who has been referred to a school psychologist for evaluation. On completion of the reading, decide whether you have sufficient information about the child to make recommendations for him as regards educational placement and treatment. If you do not feel that there is sufficient information, what additional data do you require and how will you obtain it? Detail the procedures you would follow and the methodological considerations you would make in your choice of instrumentation. Include an outline of the procedures you would follow in evaluating the success of any educational program you might recommend. Discuss the result of this exercise with other members of the group.

### Case Report

**Client:** Wadell P.
**Age:** 10-1
**Birthdate:** 2/3/65
**Date of Evaluation:** 3/10/75
**Parents:** Edward and Jane P.
327 East 800 North Boulevard
Los Angeles, California

**Referral Reason:**

Wadell P. is a black student who was referred by his teacher, Mrs. Sharon K., and Principal, Mrs. Noreen S., of A_____ School, for psychological evaluation and possible placement in a service plan for an educationally significant handicap.

**Tests Administered:**

Stanford-Binet, LM, Intelligence Test
Stanford Achievement Test
Draw-a-Man (unscored)

**Background Information:**

Wadell P. transferred to A_____ School on January 7, from a school in Los Angeles, California. His records were not available from that school at this time, but his current principal, Mrs. S___, related that Wadell's mother indicated he "had some problems" at the previous school. He was placed in the classroom of Mrs. K___, and over the past month she has voiced some concerns which necessitated this referral. Specifically, Mrs. K___ reported that Wadell does not seem to be interested in making friends or in the activities within the classroom. She describes his behavior as "impulsive" but generally "underactive, as he seems to act in slow motion." Often he "refuses to talk," and often fails to make appropriate responses in social situations (i.e., doesn't talk to other children when spoken to, doesn't join in games). She also indicated that "he is a clumsy child, or at least uncoordinated for his age. When you can get him to respond, he seems to be a litte behind in all subjects, but especially reading."

On the first day of the evaluation, Mrs. P___ was in attendance and provided the following information. Wadell is the second of four children. He has had no significant illnesses up to this time. When tested previously, his vision was believed to be 20/20 in both eyes and his hearing was "OK." Mrs. P___ stated, "I don't think there is a problem with Wadell; he's fine at home." Then was added, Mr. P___ is employed as a plumber, presently, and the family receives assistance through the Department of Family Services. Both parents reportedly completed the ninth grade. Mrs. P___ worked as a sales clerk "before getting married and becoming a housewife." Mrs. P___ reported that both parents take little interest in community affairs.

**Behavioral Observations:**

The tests were administered in two morning sessions on two consecutive days. On the first day, Wadell was brought to the testing room in the school by his teacher. When introduced to the examiner, Wadell remained quiet. At this time, Wadell was given a Draw-a-Man, which was not scored, and whose purpose was only as a rapport builder. On the second day of testing Wadell was frequently verbal with the examiner, and his speech could be characterized as "slow." During administration of the Stanford-Binet, some of the answers were impulsive, especially for the memory for digits items. Although rapport could be characterized as good, Wadell seemed quite distractable from the task at hand when the examiner would reach into the kit for the next item. Wadell remained in his seat except on two occasions. When walking, he seems to drag his feet in an uncharacteristic gait. He does not hold a pencil solidly.

**Test Results:**

The Stanford Achievement Test was administered first as a broad based assessment of Wadell's academic achievement up to this point. He was administered the Primary 1 Battery (or Intermediate II Battery) and according to the norms supplied, obtained the following scores:

| Primary I | Intermediate |
|---|---|
| a. Word Reading: one year behind | a. same |
| b. Paragraph Meaning: 10 months behind | b. same |
| c. Word Study Skills: 10 months behind | c. same |

| | |
|---|---|
| d. Vocabulary: 4 months behind | d. same |
| e. Spelling: 5 months behind | e. same |
| f. Arithmetic: 6 months behind | f. Arithmetic<br>(1) Concept Formation: 3 months behind<br>(2) Computation: 3 months behind<br>(3) Application: 5 months behind |

The Stanford-Binet was administered next as an indicator of Wadell's abstract verbal learning and problem solving ability. The test results were considered to be valid as Wadell was passively cooperative. According to the norms supplied, Wadell received an IQ=77. He had particular problems on the vocabulary, maze, and memory, and digits items. He was unable to give a reasonable answer when asked, "What would you do if you were asked by a stranger how to find the nearest service station?" (Answer: "Who," "What service station?"). Scattered throughout the other items were answers which can only be characterized as a puzzled look, as if Wadell just didn't understand the question. It is felt that Wadell had difficulty with receptive language. Questions had to be repeated; he did not engage in conversation.

**Summary**

Wadell's assessment indicated that he is performing at a low level academically, particularly in receptive and expressive oral and written language. It should be noted that when a social worker visited Mrs. P___ at home, she found that Wadell's behavior at home is in sharp variance with his behavior at school. He is capable, happy, and talkative with his peers. No one in the family is concerned about his pattern of poor performance at school.

## POSSIBLE RESPONSES TO SIMULATIONS

1. You will note that there is a significant pattern (or set of patterns) of errors made in reading this paragraph. There are numerous cases of the reversal of the letters "o" and "a" in different words. Some words were substituted for those in the test, and there was a constant omission of "the" and "an," though significantly not of "a." Words containing "i" were mispronounced. Errors in the second paragraph indicate that the student did not discriminate the word, "stove." He/she seemed to expect that every reference would be to "bed." However, he/she did make a correction on both occasions. Any remediation should focus on these deficits. Your criterion-referenced test should be constructed to monitor progress on these areas. When constructing the test, remember to include criteria for mastery and to rationalize your choice.

2. Information presented in this case study is not adequate for you to make an appropriate decision about educational placement and treatment for Wadell. There seems to be a discrepancy between Wadell's behavior at home and at school. You would need more information about such variables as the teacher's expectations of Wadell, demands made on him by the classroom environment, and the nature of reinforcement methods employed in the classroom.

   The testing instruments used should be evaluated to determine if they were appropriate, particularly in terms of language, concepts tested, relationship to curriculum, etc.

   You would recommend that further assessment be conducted. A thorough assessment of language, both receptive and expressive, is necessary. You would also conduct a profile analysis to determine any reliable differences between subtests on the SAT. You would recommend diagnostic assessment in deficit areas, and further criterion-referenced information. You would, of course, outline the considerations you would make in selecting the instruments, especially those concerning reliability and validity.

   Following the assessment process, you would examine the curriculum in which Wadell was placed to determine which deficit areas are the results of inadequate learning experiences and which are the results of learning problems.

## KEY TO PRETEST

**Short Answer Items**

1. To make educational decisions about the performance of students.
2. Testing is only one method of assessing a child. Assessment is a flexible, continuous process leading to an ongoing program which may be modified in the interests of the student's life situation and of a reduction of his/her current specific difficulties.
3. a. Selection and administration of tests in a nondiscriminatory manner.
   b. Provision of test materials and procedures in the child's native language or mode of communication.
   c. Prohibition of the use of a single test or procedure as the sole criterion for determining educational placement.
4. a. Child's files.
   b. Information from significant other in the child's life.
   c. Observation of the child.
   d. Norm-referenced tests.
   e. Criterion-referenced tests.
5. a. Current life circumstances.
   b. Developmental history.
   c. Extra-personal factors.
   d. Situational factors.
6. a. Referral.
   b. Screening.
   c. Evaluation of screening and planning interventions.

    d. EValuation of intervention effects.
7. a. Observer bias.
    b. Awareness by the student that he/she is being assessed.
    c. Inconsistency between interviewers.
    d. Test invalidity and/or unreliability.
8. A student's scores on an achievement test reflect pupil ability as compared to others of comparable age level and the instructional effectiveness of an educational program.
9. a. Use of the test for purposes other than those for which it was designed.
    b. Use of a test for individual purposes when designed for use with a group.
10. a. To provide recommendations for instructional and behavioral management.
    b. To provide recommendations for educational placement.
11. Norm-referenced tests evaluate the child's performance by comparison with that of other children. Criterion-referenced tests evaluate the child in terms of some learning task at some absolute standard that is independent of other children's performances.
12. A domain-referenced test is one in which the emphasis is placed on precisely identifying skill areas (domains). The test items of a domain-referenced test are selected so that one can be certain that the child who meets criteria on the items which represent a domain (such as long division of decimal fractions using single digit divisors) would be able to meet all the other examples of a domain when they were encountered.
13. a. Criterion should be based on subject matter.
    b. The range should be 80-100 percent.
    c. The criterion should relate to the nature of the test.
    d. Performance of others.
    e. Expert decision on minimal competence.
14. Nature of the change in the student and impact of the intervention.
15. The degree to which tests enable educators to accurately describe a treatment effect.
16. a. Parental permission is necessary before initiation of diagnosis or evaluation and for placement beyond a regular education classroom.
    b. Multidisciplinary evaluation.
    c. Due process or procedural safeguards to ensure the educational rights of the child.

**True/False Answers**

1. F
2. F
3. T
4. F
5. F
6. T
7. T
8. F
9. F
10. F
11. T

## KEY TO POSTTEST

1. d
2. d
3. a
4. d
5. a
6. c
7. d
8. a
9. c
10. c
11. b
12. b
13. c
14. c
15. b
16. c
17. b
18. a
19. b

20. a. Validity of the test for the purpose and the populations.
    b. Information about technical adequacy of test
        (1) administration
        (2) standardization/norming groups.
    c. Knowledge of the physical and mental characteristics of the child being tested.
    d. Knowledge of curriculum in which child is currently placed.
    e. Reliability coefficients (for test and subtests).
21. Answers should include: Characteristics of geographic region, cultural ethnic background, age of student (C.A., and M.A.), nature of disability and its possible effect on nature and rate of response; any medication prescribed for the child and its influence on the nature and rate of response.
22. a. What is the reliability of the data supporting a change?
    b. What is the direction of the change?
23. a. The lack of relationship between test and treatment.
    b. The quality of treatment.
    c. The lack of relationship between treatment and curriculum.

24. Administrator's Needs
    Student program placement.
    Evaluation of educational progress.
    Appraisal of effectiveness of specific curricula and programs.
    Identification of students with specific disabilities for funding purposes.

    Teacher's Needs
    Instructionally relevant information.
    Precise knowledge of students' specific skills and deficits.
    Any other information that will facilitate effective intervention.

**Practical Implications**

Nature of assessment instruments (norm-referenced, criterion-referenced, diagnostic rather than achievement test, relationship between test and curriculum sequence).

## REFERENCES

Adelman, H. S., & Taylor, L. Initial psychoeducational assessment and related consultation. *Learning Disability Quarterly*, 1979, 2, 52-64.

Burton, N. W. Societal standards. *Journal of Educational Measurement*, 1978, 15(4), 263-271.

Clarke, A. M., & Clarke, A. D. B. *Mental deficiency: The changing outlook*. New York: Free Press, 1975.

*Diagnosis: An instructional aide—mathematics, levels A and B*. Chicago: Science Research Associates, 1980.

Donlon, T. F. Referencing tests scores: Introductory concepts. In W. Hively, and M. C. Reynolds (Eds.), *Domain-referenced testing in special education*. Minneapolis, Minnesota: Special Education Leadership Training Institute, University of Minnesota, 1975.

Gallery, M. & Hofmeister, A. M. A method for assessing the treatment validity of tests in special education. *The Exceptional Child*, 1978, 25(2), 105-113.

Glaser, R. Instruction, Technology, and the measurement of learning outcomes: Some questions. *American Psychologist*, 1963, 18(8), 519-521.

Hickey, M. E., & Hoffman, D. H. Diagnosis and prescription in education. *Educational Technology*, 1973, 13(10), 35-37.

Hively, W. Introduction to domain-referenced testing. *Educational Technology*, 1974, 14(6), 5-10.

Hofmeister, A. M. Testing and treatment validity. *Journal of Learning Disabilities*, 1979, 12(3), 73-75.

Hofmeister, A. M. *Diagnostic capital letters and punctuation tests*. Logan, UT: Exceptional Child Center, Outreach and Development Division, Utah State University, 1972.

Housden, J. L. & LeGear, L. An emerging model: Criterion-referenced evaluation. *Thrust*, April, 1973.

Kamin, L. J. Social and legal consequences of IQ tests as classification instruments: Some warnings from our past. *Journal of School Psychology*, 1975, 13, 317-323.

Lien, A. J. *Measurement and evaluation of learning*. Dubuque, IA: W. C. Brown, 1971.

McIntyre, K. H. Adaptive behavior. In A. L. Pasanella, T. B. Sales, & C. B. Volkmor (Eds.), *Cross-cultural educational assessment skills workshop: Training processes and materials*. Los Angeles: California Regional Resource Center, 1978.

Mischel, W. *Personality and assessment*. New York: Wiley, 1968.

Phye, G. School psychologists as consultants in the evaluation of learning and intervention outcomes. In G. D. Phye & D. G. Reschly (Eds.), *School psychology: Perspectives and issues*, New York: Academic Press, 1979.

Reynolds, M. C. Trends in special education: Implications for measurement. In W. Hively and M. C. Reynolds (Eds.), *Domain-referenced testing in special education*. Minneapolis, Minnesota: Special Education Leadership Training Institute, University of Minnesota, 1975.

Salvia, J. & Ysseldyke, J. E. *Assessment in special and remedial education*. Boston: Houghton-Mifflin, 1978.

Shertzer, B., & Linden, J. D. *Fundamentals of individual appraisal: Assessment techniques for counselors*. Boston: Houghton-Mifflin Co., 1979.

Thurlow, M. L. & Ysseldyke, J. E. Current assessment and decision-making practices in model LD programs. *Learning Disability Quarterly*, 1979, 2(4), 15-24.

Wechsler, D. *Manual for the Wechsler Intelligence Scale for Children, Revised*. New York: Psychological Corporation, 1974.

Ysseldyke, J. E., & Algozzine, B. Perspectives on assessment of learning disabled students. *Learning Disability Quarterly*, 1979, 2(4), 3-13.

## RECOMMENDED READING

Adelman, H. S. & Taylor, L. Initial psychoeducational assessment and related consultation. *Learning Disability Quarterly*, 1979, 2, 52-64.

**Abstract:** Criticism of current approaches to assessment indicates the need for effective alternatives. Given a general understanding of the criticism, the authors describe the basis for an alternative approach to initial assessment activity. Specifically, the discussion focuses on (1) a conceptualization of initial assessment and consultation; (2) the problem-solving paradigm as a framework for guiding this activity; and (3) a description of procedures and initial data from a demonstration program.

Becker, W., & Engelmann, S. *Teaching 3: Evaluation of Instruction*. Chicago: Science Research Associates, 1976.

**Abstract:** This book elaborates and exemplifies many of the constructs presented in this section on educational assessment. Included in this book are the following topics: kinds of tests; understanding test scores; constructing instructional-program-based tests; evaluating instructional-program-based tests; approaches to monitoring student progress; outcome evaluation with criterion-referenced tests; norm-referenced achievement tests; interpreting norm-referenced test scores; and, who can be taught.

Gallery, M. & Hofmeister, A. M. A method for assessing the treatment validity of tests in special education. *The Exceptional Child*, 1978, 25(2), 105-113.

Hofmeister, A. M. Testing and treatment validity. *Journal of Learning Disabilities*, 1979, 12(3), 73-75.

**Abstract:** The above two articles include specific recommendations to the psychologist who is concerned about selecting tests that will have instructional relevance. A detailed frame of reference listing specific criteria for test selection is provided.

*Learning Disability Quarterly*. 1979, 2(4).

**Abstract:** This whole issue highlights the assessment of learning disabilities. Three articles that are especially applicable have been included in the Annotated Bibliography.

Salvia, J., & Ysseldyke, J. E. *Assessment in special and remedial education.* Boston: Houghton-Mifflin, 1978.

**Abstract:** This text is designed primarily for teachers in special and remedial education, but also for the support system of special education students, including counselors, educational administrators, school psychologists, and social workers. No prior knowledge of measurement and statistical concepts is assumed.

Parts 1 and 2 provide a general overview of an orientation to assessment. Part 3 provides detailed discussions and assessment of achievement, intelligence, perceptual-motor skills, sensory functioning, language, personality, adaptive behavior, and readiness. Part 4 is integrative and deals with the application of assessment practices in special and remedial education. Part 4 is felt to be particularly applicable to this module.

Thurlow, M. L., & Ysseldyke, J. E. Current assessment and decision-making practices in model LD programs. *Learning Disability Quarterly,* 1979, 2(4), 15-24.

**Abstract:** Data from questionnaires completed by forty-four Child Service Demonstration Centers were analyzed. Information was provided on the number of children served, the LD definition used, the kinds of assessment data collected, and the purpose for which they were used; the specific assessment devices used to collect data, and the purpose for which they were used; the typical composition of the placement team, and the major sequential steps in the assessment/decision-making process. Results suggested that assessment and decision-making in the field of learning disabilities are characterized by variability and inconsistency. The implications of the findings for current assessment practices are summarized.

Ysseldyke, J. E., & Algozzine, B. Perspectives on assessment of learning disabled students. *Learning Disability Quarterly,* 1979, 2(4), 3-13.

**Abstract:** Current critical issues in assessment of learning disabled students are described with special emphasis on logical fallacies in the assessment process. New directions in assessment are specified and discussed.

# NONBIASED ASSESSMENT AND MILDLY RETARDED STUDENTS

Daniel Reschly
*Iowa State University*

# Table of Contents

Introduction .......... 97

Objectives .......... 97

Pretest .......... 99

Mental Retardation: Basic Concepts .......... 101
  Definitions and Classification Criteria .......... 101
  Mild Mental Retardation: "A Continuing Dilemma" .......... 101

Appropriate Assessment: Mild Mental Retardation .......... 102
  Multifactored Assessment .......... 102

Intelligence .......... 102
  Selection of Intelligence Tests .......... 103
  Proper Interpretation of Intelligence Test Results .......... 105
  Summary .......... 106

Adaptive Behavior .......... 106
  Adaptive Behavior and Definitions of Mental Retardation .......... 107
  Conceptions of Adaptive Behavior .......... 107
  Analysis of Conceptions of Adaptive Behavior .......... 107
  Assessment of Adaptive Behavior .......... 108
  Review of Adaptive Behavior Measures for Mildly Retarded Students .......... 109

Nonbiased Classification/Placement in Mild Mental Retardation .......... 111
  Analysis of Data on Placement Bias .......... 111
  Bias in Intelligence Tests .......... 112
  Perspective on Nonbiased Assessment .......... 112
  Combining Intelligence and Adaptive Behavior Data .......... 113

Summary .......... 115

Simulations .......... 115

Pretest Key .......... 119

References .......... 119

Recommended Reading .......... 121

## INTRODUCTION

School psychologists and special educators have been concerned with the classification and special class placement of mildly retarded children and youth throughout the twentieth century. Much of the history of special education and school psychology has been dominated by efforts to identify and provide special class programs for student who are so classified. Although the scope and nature of these interrelated fields have expanded and changed in recent years, concern with the mildly retarded population continues.

Mild mental retardation is not understood well. Questions still arise over definition, etiology, and psychological and educational needs. One of the most intense debates focuses on the overrepresentation of minority children and youth in programs for the mildly retarded. The overrepresentation issue, called "placement bias" by many concerned persons, was the basis for litigation in several localities during the decade of the 1970s. The explicit allegation in most of the placement litigation was that the overrepresentation of minorities was caused by inappropriate assessment, particularly by bias in measures of ability. The litigation was followed by federal and state legislation that has a number of purposes, one of the most important being the elimination of discrimination against ethnic or racial minorities in special education classification-placement decisions.

This section of the module provides information on and simulation activities for the appropriate assessment of mildly retarded children and youth. Special attention is devoted to some issues in the assessment of minorities and concepts of bias in assessment. The following general objectives are established for this section.

### Objectives

1. Describe the major definitions of mental retardation, identify similarities and differences in definitions, and describe characteristics of mildly retarded students.
2. Describe the sequence, kind of data, source of data, and appropriate uses of the components of a multifactored assessment. For each step, specify how the data would be used to assess a child who might be mildly retarded.
3. Describe the use of intelligence test information in classification/placement decisions for students with mild mental retardation.
4. Describe the use of adaptive behavior information in classification/placement decisions for students with mild mental retardation.
5. Define the concept of placement bias and analyze data suggesting placement bias.
6. Describe different definitions of bias in tests and bias in assessment, and relate the research results pertinent to each definition.
7. Discuss different ways to combine intelligence and adaptive behavior information in classification/placement and programming decisions.
8. Given extensive data on a child, including intelligence and adaptive behavior information, suggest and justify (a) an appropriate classification decision and (b) different educational alternatives for meeting the student's needs.

The content included here is not sufficient in depth or scope to provide all the competencies which are needed by persons who participate in classification, placement, and programming decisions for the mildly retarded. The minimum competencies for diagnostic personnel in mental retardation as described by Cegalka (1978) are reproduced in Table 1. Diagnostic personnel are encouraged to use the information in Table 1 to examine their competencies and to identify needs for continuing education.

### Table 1

**Recommended Areas of Study and Competencies for Professionals Who Are Responsible for Classifying Mentally Retarded Persons**

| Recommended Areas of Study | Expected Competencies |
|---|---|
| | **Individual Assessment** |
| Nature of intelligence<br>Developmental milestones<br>Behavioral observation<br>Test selection<br>Administration of individual scales, tests, or other measures of intelligence, social adaptation, perceptual motor performance, academic achievement, language development, or personality<br>Scoring and test profile with developmental, behavioral, and other data<br>Relation of measurement to individual programming | Is familiar with and can critically evaluate a wide range of intelligence and other assessment instruments<br>Selects appropriate battery of tests or scales to answer specific questions for various ages, functioning levels, and disability groups<br>Makes reliable and pertinent behavioral observations and can report observations clearly<br>Can assess the functioning level of infants and of mute, blind, deaf, or physically handicapped persons with various levels of intellectual functioning.<br>Can integrate measurement results, developmental data, and information from other disciplines into a meaningful report that is useful for program development |

| | |
|---|---|
| Written and oral communication of results<br>Consultation and follow up<br>Cross cultural studies, including the effects of linguistic background on test results | Reports results of assessment clearly in written and oral form to parents and professionals<br>Is aware of the effects of situational factors and cultural background on test results<br>Can accurately administer scales, tests, or other measures under a variety of working conditions<br>Can formulate a plan for monitoring the efficacy of recommendations derived from test results<br>Is familiar with subcultural and ethnic group vocabularies |

**Individual Differences**

| | |
|---|---|
| Theories of development<br>Developmental sequence<br>Learning and development<br>Maturation and test behavior | Has knowledge of major theories of development<br>Knows the possible effects of maturation upon test behavior at different ages<br>Is familiar with and can readily identify developmental sequences, milestones, and behaviors that are found through the life range |

**The Exceptional Child**

| | |
|---|---|
| Survey of disability groups<br>General and specific learning disabilities<br>Differential assessment<br>Educational alternatives | Is knowledgeable about the characteristics, differential assessment, and learning needs of a variety of disability groups and can relate learning needs to specific educational alternatives |

**Learning and Remediation**

| | |
|---|---|
| Application of learning and educational theories of remediation<br>Interdisciplinary planning and evaluation of education outcomes<br>Development of individualized program plans including curricula<br>Techniques of supervision and consultation | Can use a given learning or educational theory to devise specific educational plans<br>Can write behavioral objectives, incorporating appropriate criterion measures and accountabilities<br>Can function as a contributing member of an interdisciplinary team for the purposes of individual assessment and program development<br>Is comfortable in working on a counseling or consulting basis with parents or professionals, either individually or in groups |

**Mental Retardation**

| | |
|---|---|
| Historical trends and issues in the field<br>Etiology, incidence, and prevalence<br>Terminology and classification<br>Adaptive behavior levels<br>Assessment of mental retardation<br>Program needs and alternatives<br>The family and mental retardation<br>Resources and future directions in the field | Is knowledgeable about historical trends and modern issues in the field<br>Uses appropriate terminology and is aware of possible destructive implications of labeling and self defeating expectancies<br>Is aware of program alternatives for each level of mental retardation throughout the life range and can describe a desirable continuum of services<br>Can differentiate mental retardation from mental illness or other conditions that may lower functioning level<br>Can select an appropriate battery of tests regardless of the retarded person's age, functioning level and multiple handicaps<br>Uses test results and other pertinent information to devise educational plans for all ages and levels of mental retardation<br>Can identify major information resources and referral agencies in the field |

**Supervised Experience**

Experience in a variety of program settings

Supervised assessment of mentally retarded persons representing all ages and levels of functioning

Participation in multidisciplinary team assessment and decision making

Development of remediation training goals

Report writing

Parent conferences

Monitoring and reasssessment of recommendations derived from test results

Is knowledgeable about assessment procedures and related programming strategies used in a variety of settings including diagnostic centers, early education programs, day training centers, special education, vocational training and placement

Uses an appropriate battery of test instruments to assess the functioning level of mentally retarded persons with different intellectual levels and concomitant handicaps

Communicates assessment information to persons in a useful manner

Integrates test results with information from other disciplines to derive recommendations for remediation or training.

Makes realistic plans for monitoring the implementation of recommendations

---

Source: Reprinted from W. Cegalka, Competencies of persons responsible for classification of mentally retarded persons. *Exceptional Children*, 1978, 45, 26-31, Table 1.

Sources of additional information that are fundamental to nonbiased assessment with the mildly retarded are listed at the end of this section.

## PRETEST

### Nonbiased Assessment and Mildly Retarded Students

I. Mental Retardation: Basic Concepts

1. What are the two major sources of classification criteria in mental retardation?
   a. APA Diagnostic Manual and State Education Codes
   b. PL 94-142 Rules and Regulations and AAMD *Manual on Terminology and Classification*
   c. Diagnostic and Statistical Manual (DSMIII) and PL 94-142 Rules and Regulations
   d. State Education Codes and AAMD *Manual on Terminology and Classification*
2. The most common educational term for mild mental retardation is
   a. Borderline mental retardation
   b. Educable mental retardation
   c. Trainable mental retardation
   d. Slow learner
   e. Educationally handicapped
3. True-False. According to the AAMD *Manual*, mental retardation is usually permanent.
4. True-False. According to the AAMD *Manual*, mental retardation may be due to psychosocial, psychogenic, or biological causes.
5. Mild retardation differs from moderate, severe, and profound retardation in terms of
   a. permanence of the condition
   b. etiology
   c. comprehensiveness
   d. socio-economic status
   e. all of the above

II. Appropriate Assessment: Mild Mental Retardation

A. Multifactored Assessment

6. True-False. The first step in a multifactor assessment process for a child who may be classified as mildly retarded is intellectual assessment.

B. Intelligence

7. In the special education placement *litigation*, the overrepresentation of minorities was usually attributed to
   a. bias in IQ tests
   b. differential referral rates for minority and nonminority students
   c. Inadequate educational alternatives for low-achieving minority students
   d. all of the above
8. True-False. Brief IQ measures generally are acceptable for classification decisions in mental retardation if the brief measure has a high correlation with conventional IQ tests.
9. Which of the following is not an underlying assumption in intellectual assessment?
   a. Maximum performance
   b. Skilled examiner

   c. Comparable acculturation
   d. Comparison to local norms

10. True-False. If two tests correlate perfectly, we can assume that an individual's scores on the tests are comparable.

11. The average amount of scatter (subtest variation) on the WISC-R for normal children is about
   a. 7 scaled score points
   b. 5 scaled score points
   c. 3 scaled score points
   d. 0 to 1 scaled score points

12. True-False. A flat profile for someone in the borderline range of intelligence is indicative of an underlying process of mild mental retardation.

13. Which of the following is *not* a common myth about the meaning of IQ tests results?
   a. Intelligence is unitary
   b. Intelligence is fixed
   c. Intelligence is predetermined
   d. Intelligence is culturally determined

14. A good synonym to replace the term intelligence in describing what IQ tests measure is
   a. innate potential
   b. ability to adapt
   c. academic aptitude
   d. intellectual potential

### C. Adaptive Behavior

15. True-False. Adaptive behavior and intelligence mean nearly the same thing.

16. The emphasis on adaptive behavior in recent litigation and legislation reflects the underlying purpose of
   a. reducing bias in classification/placement decisions
   b. promoting normalization through identifying and remediating deficit behaviors among the severely retarded
   c. differential diagnosis of emotionally disturbed, learning disabled, and mildly retarded students
   d. all of the above

17. True-False. Nearly every state lists some criteria to be used in assessing adaptive behavior.

18. The two common features among different conceptions of adaptive behavior are
   a. maturation and learning
   b. cross-cultural expectations and general trait views
   c. age appropriate criteria and cultural context
   d. academic achievement and intelligence
   e. all of the above

19. True-False. According to the AAMD *Manual*, school achievement is an important domain of adaptive behavior during the school-age years.

20. Most currently available measures of adaptive behavior, such as the AAMD Adaptive Behavior Scale, originally were designed for what purpose?
   a. Intervention/Programming with moderately, severely, and profoundly retarded students.
   b. Classification/Placement with mildly retarded students.
   c. Identification of needs for social skills training with normal children.
   d. Differential diagnosis of mental retardation at very young ages.

21. Which of the following domains is *NOT* included in the Adaptive Behavior Inventory for Children?
   a. Earner/Consumer roles.
   b. Peer Group relations.
   c. Academic/School related activities.
   d. Self-maintenance.
   e. Community roles.

22. The norms for the ABIC and AAMD-Public School measures are based on samples of children from
   a. western states
   b. throughout the United States
   c. all regions except the South
   d. California

### III. Nonbiased Classification/Placement in Mild Mental Retardation

23. Approximately what proportion of black children were placed in programs for the mildly retarded in the Larry P. case?
   a. 50-66%
   b. 33-50%
   c. 10-33%
   d. 5-10%
   e. 1-3%

24. **True-False.** Placement bias is specified as the definition of discrimination in testing and placement in the PL 94-142 Rules and Regulations.

25. **True-False.** Intelligence tests are less reliable and less valid for making predictions when used with minorities.

26. **True-False.** Placement bias and item content bias are two widely agreed upon examples of bias in assessment.

27. Resource rather than special class programs are more appropriate for most of the students who are
    a. comprehensively retarded
    b. quasi-retarded
    c. moderately retarded
    d. multiply handicapped

## Mental Retardation: Basic Concepts

### Definitions and Classification Criteria

Terminology in and classification criteria for mental retardation have evolved throughout this century. Two major sources of terminology and classification criteria are crucial for diagnostic personnel in the public schools: the *Manual on Terminology and Classification*, published by the American Association on Mental Deficiency (AAMD), the major professional organization for mental retardation; and state education codes.

The AAMD Manual is revised periodically; the latest revisions were in 1961, 1973, and 1977 (Grossman, 1973, 1977; Heber, 1961). The 1973 and 1977 revisions are virtually identical. The AAMD *Manual* significantly influences other definitions and classification criteria in mental retardation, and the influence of the 1961 and 1973 versions often are reflected in state education codes. The latter usually provide the definition and classification criteria for mental retardation which public school diagnostic personnel must apply. Although the AAMD System may be important, decisions in the public schools must be based on the *state* definition and criteria for mental retardation. Knowledge of the current state code, usually published as part of the state special education rules and regulations, is necessary for diagnostic personnel.

The basic definition of mental retardation in the 1973 and 1977 revisions of the AAMD *Manual* is also used in the PL 94-142 Rules and Regulations. It is as follows:

> Mental retardation refers to significantly subaverage general intellectual functioning existing concurrently with deficits in adaptive behavior, and manifested during the developmental period. (Grossman, 1977, p. 11)

MacMillan (1977) and Robinson and Robinson (1976) provide thorough analyses of the AAMD classification system in mental retardation. Some of the most important characteristics of the system follow:

1. **Bi-Dimensional.** The individual must exhibit deficits in *both* intelligence and adaptive behavior in order for the classification of mental retardation to be appropriate.
2. **Developmental.** The deficits in intelligence and adaptive behavior must appear during the developmental period (i.e., between birth and 18).
3. **Current Status.** "Mental retardation is descriptive of current behavior and does not necessarily imply prognosis" (Grossman, 1977, p. 11).
4. **Etiology.** Etiology of mental retardation is not specified in the definition. The etiology may be psychosocial, psychogenic, or biological.
5. **Continuum.** All types and levels of mental retardation are implicitly organized on the same continuum ranging from mild to profound.
6. **Levels.** The level or severity of mental retardation is specified by standard deviation (s.d.) cut-off points for IQ tests scores:
   Mild (Educable): -2 s.d. IQ Range of about 55-69
   Moderate (Trainable): -3 s.d. IQ Range of about 40-54
   Severe: -4 s.d. IQ Range of about 25-39
   Profound: -5 s.d. IQ Range of about 24 and below
7. **Adaptive Behavior.** The criteria for adaptive behavior depend on the age of the person. (See subsequent discussion.)

Two significant changes were made in the 1973 revision of the AAMD *Manual*. (a) The "borderline" level of mental retardation (i.e., IQs between 70 and 84) was deleted. (b) Increased emphasis was placed on adaptive behavior (see later section).

Most states use the current AAMD definition or a variation thereof; nevertheless, terminology and classification criteria vary from state to state. For example, the terms "mental handicap" or "mental disability" may be used to refer to mental retardation. In many states the AAMD levels of mental retardation are modified by the adjectives "educable" and "trainable" rather than "mild" and "moderate." Classification criteria in state definitions also vary in terms of IQ score cut-off points, conception of adaptive behavior, and amount of emphasis on adaptive behavior. A number of states continue to use a cut-off score higher than 70 to define the highest level of mental retardation.

---

*Use Simulation 1 here.*

*All Simulations appear at the end of the narrative text.*

---

### Mild mental retardation: "A continuing dilemma"

Mild retardation has been a principal concern in recent litigation on the overrepresentation of minorities in special education. The definition of mild mental retardation, of course, varies from state to state. Mild mental retardation in an educational setting traditionally has meant a child who is (a) referred for academic and/or behavioral problems, (b) exhibits educational achievement significantly below age or grade expectations, and (c) obtains an IQ score well below average, with cut-off scores ranging from about 55-69 to as high as 60-84. The meaning of, criteria for, and method to assess adaptive behavior vary considerably (see subsequent discussions).

Children who exhibit these three symptoms often are classified as mentally retarded and may be placed in different types of special education programs. However, the characteristics of children with mild or educable mental retardation often are confused with those of the more severe levels of mental retardation. In fact, the differences between the mild and more severe levels of mental retardation are sufficiently large to raise questions about the appropriateness of using the same continuum in educational settings for all forms of mental retardation. In addition to the obvious characteristics of higher IQ score, some more salient differences between mild and lower levels of mental retardation are as follows:

1. **Etiology.** Consensus on the etiology of mild mental

retardation has not been achieved. The most common etiology of mild mental retardation given in the 1977 AAMD *Manual* is the category of "psychosocial" (in the 1961 version, the term used was "cultural-familial"). In contrast to the moderate, severe, and profound levels of mental retardation, there is no evidence of biological anomaly in the vast majority of cases of mild mental retardation. Whether mild retardation represents the lower tail of the normal distribution of genetic potential, the complex effects of extreme poverty, or a specific defect in cognitive functioning was called "a continuing dilemma" by Zigler (1967). Other explanations include Mercer's (1973) view that most diagnosed cases represent an "acquired status" attributable to cultural differences.

2. **Age.** Mild mental retardation rarely is diagnosed prior to age 5 or 6, and the highest prevalence usually is found in late childhood or early adolescence.
3. **Social System.** The behavioral deficits of children who are classified as mildly retarded usually are restricted to the public school setting. Performance in other settings usually is regarded as normal by significant others, such as parents, siblings, and adults in the larger community.
4. **SES.** There is a strong association between the prevalence of mild mental retardation and socio-economic status (SES) for all sociocultural groups. Low SES children, regardless of race or ethnicity, are more likely to be classified as mildly retarded.
5. **Ethnic-Racial Group.** The prevalence of mild mental retardation is higher in specific ethnic or racial groups who, in turn, are much more likely to be of low SES. Whether SES accounts for the overrepresentation of specific groups or additional factors of bias are responsible is the subject of continuing debate. However, the data on overrepresentation often are misunderstood (see subsequent discussion).
6. **Adult Adjustment.** The vast majority of persons classified as mildly retarded during the school-age years become self-supporting, independent functioning adults (Baller, Charles, & Miller, 1967). However, most are in low-status occupations with somewhat marginal economic and social adjustment.

In view of these characteristics, the popular stereotype of mental retardation as incompetence that is comprehensive, permanent, and biologically based is particularly troublesome. In fact, the vast majority of mentally retarded persons are mildly retarded. Most persons who are diagnosed as mildly retarded do not have identifiable physiological anomalies, they function within normal limits as adults, and their mental retardation is specific to the public school setting. A desirable step in relation to the stereotype and the differences between mild and lower levels of mental retardation would be to change the classification system. Terminology other than mental retardation might be more appropriate for persons now classified as mildly retarded. However, until the classification system is refined, we must devote considerable effort to communicating to teachers, parents, and students that mild mental retardation is not comprehensive, permanent, or biologically based.

---

*Use Simulation 2 here.*

---

## Appropriate Assessment: Mild Mental Retardation

### Multifactored Assessment

In this section, emphasis is given to the dimensions of intelligence and adaptive behavior in the diagnostic criteria for mental retardation. Yet, a complete multifactored assessment (Tucker, 1977) should be conducted before any child is classified as mildly retarded. Furthermore, all the "Protection in Evaluation Procedures" (see Appendix 1) and "Due Process Procedural Safeguards" (see section of this module by D. Bersoff, "Nonbiased Assessment: Legal Principles") of the PL 94-142 Rules and Regulations should be followed.

The conceptual basis and major components associated with a multifactor assessment are presented elsewhere (see section of this module by T. Oakland, "Nonbiased Assessment: Basic Considerations"). Other sections of this module are devoted to specific components of multifactored assessment. An adaptation of Tucker's (1977) description of multifactored assessment is presented in Table 2. In most cases the assessment activities should proceed in the sequence suggested in the table. The purpose of this sequence is to ensure the collection and consideration of a variety of information and, implicitly, to ensure that intelligence test data do not dominate the decision-making process. Much of the information gathered in the earlier steps is useful for clarifying the nature of the problem, developing interventions in the regular classroom, and (if the process continues to later steps) selecting appropriate individual intelligence tests, and interpreting intelligence test results.

---

*Use Simulation 3 here.*

---

## Intelligence

> Mental Retardation refers to significantly subaverage general intellectual functioning existing concurrently with deficits in adaptive behavior. . . . (Grossman, 1977, p. 11)
>
> General intellectual functioning is defined as the results obtained by assessment with one or more of the individually administered general intelligence tests developed for the purpose. (Grossman, 1977, p. 11)
>
> While IQ is only one of the factors that are evaluated in arriving at a diagnosis of mental retardation, it is a crucial one. (Sattler, 1974, p. 308)

Intelligence for centuries has been and continues to be a fundamental concept in the efforts to describe and understand mental retardation. Intelligence is a prominent feature in all major definitions of mental retardation. Attempts to diagnose mental retardation without considering intelligence are virtually unheard of.

Despite the central role of intelligence in the concept and diagnosis of mental retardation, the entire enterprise of intellectual assessment has been the focus of intense debate and criticism in recent years. Much of the controversy focuses on the use of intelligence test results as part of the basis for classifying and placing children in special education programs. Overrepresentation of minority students in programs for the mildly retarded has been and continues to be a paramount issue in this debate.

Intelligence test results have been viewed frequently as the primary or even sole reason for the overrepresentation of minorities in programs for mildly retarded children and youth. This rather simplistic and probably inaccurate notion is apparent in a number of sources of which the most notable is the opinion in the *Larry P. vs. Riles* case (1979). The assumptions that intelligence tests are biased and that intelligence test results were the primary, if not the sole, basis for classification were only a part of the complex issues involved in the placement litigation (Reschly, 1979).

It is unclear whether overreliance on intelligence test

## Table 2
### Sequential Outline of Multifactored Assessment

| Type of Data | Purpose | Data Source |
|---|---|---|
| 1. Observational | Determine degree of deviance from expectations | Natural setting; teachers and others who have direct contact with child |
| 2. Other data available | Disconfirm or corroborate the nature and and extent of problem; consider other factors such as medical or temporary, situational conditions | Cumulative records, previous evaluations of child, records and/or test scores from previous years |
| 3. Language dominance | Determine child's primary language, assess whether the school-related problem is due to language differences, and decide appropriate language for administration of other instruments | Tests of primary language and interviews with parents |
| 4. Educational | Determine level, pattern, and strengths and weaknesses in academic areas; confirm or disconfirm discrepancies established in previous steps | Formal and informal achievement measures administered by trained diagnosticians |
| 5. Sensory-motor (optional) | Determine whether deficits in process areas are possible explanations for the learning problem | Usually formal tests administered by trained diagnosticians |
| 6. Adaptive behavior | Determine social competence outside of school in relation to age and cultural expectations | Usually structured interview with parents by trained diagnosticians |
| 7. Medical and/or developmental | Determine if medical, sensory, or health factors are related to the onset or treatment of the learning problem | Developmental history, medical records, and if indicated, physical examination |
| 8. Personality assessment, including self-report | Determine degree of emotional involvement (if any) and obtain student's perception of problem | Interviews and formal or informal devices administered and interpreted by appropriately trained persons |
| 9. Intellectual | Predict level of achievement, and with other data, classify if appropriate | Formal tests measuring verbal and nonverbal abilities administered by trained specialists |

Adapted from J. Tucker, Operationalizing the diagnostic-intervention process. In T. Oakland (Ed.), *Psychological and educational assessment of minority children*. New York: Brunner/Mazel, 1977.

results is primarily to blame for the overrepresentation of minorities in special classes. MacMillan (1977) pointed out that intelligence testing is not and never has been the first step in special education classification decisions. Initial teacher referral and prior academic achievement, potentially at least, are just as important influences on the pattern of overrepresentation. Moreover, intelligence test results actually may protect children, particularly minority children, from erroneous classification decisions (Ashurst & Meyers, 1973). If intelligence tests were not used and decisions were based on teacher referral, classroom academic performance, and the results of standardized achievement tests, the overrepresentation of minorities might be even larger.

As noted previously, intellectual assessment should be only one part of the multifactor assessment process. Full attention to the scope and sequence of a multifactor assessment has not occurred in all school districts. The information from earlier steps in the multifactor assessment process is needed to select an appropriate intelligence test and to properly interpret the results.

**Selection of Intelligence Tests**

A number of general principles are accepted almost universally regarding the selection of a test to assess intelligence when a diagnosis of mental retardation is under consideration.

The test should be as reliable and valid as possible.

The norms for the test should be based on a representative sample of the population of which the child is a part.

The response format of the test should be appropriate to the child.

The examiner should be skilled in the administration of the test, knowledgeable about normal and abnormal patterns of development, capable of observing qualitative features of test performance, and proficient in interpreting the results and developing intervention programs.

Underlying the entire enterprise is the crucial assumption of *maximum* performance (Anastasi, 1976), that is, that the child is exhibiting his/her best effort and performance.

This list of factors which are relevant to the selection of tests and examiner competencies could go on and on. Unfortunately, many of the principles are violated even though almost universal agreement exists regarding their importance. Two illustrations of major violations follow. (See Sattler, 1974, for a comprehensive overview of intellectual assessment, and MacMillan, 1977, and Robinson & Robinson, 1976, for discussion of the intellectual assessment of mentally retarded students.)

**Illustration 1. Brief screening measures.** In some instances, brief screening devices or even group tests have been used to measure a child's intellectual dimension when a classification of mild mental retardation is possible. Although the practice is probably infrequent, the fact that it occurs at all is reason for concern. The use of brief measures or group tests clearly is inconsistent with recommendations in the AAMD *Manual*.

> The measures should be individually administered standardized general intelligence tests in every case where the individual under consideration can respond to such tests. Under no circumstances should an individual be classified as mentally retarded in intellectual functioning on the basis of a group test or on the basis of one of the "quick" tests; group tests may produce low scores for any number of reasons, none of which are evident in the test pages. Individual intelligence tests should be administered for the purposes of determining retardation in intellectual functioning only if the examiner is specifically trained and has had supervised experience in use of tests, test construction, and with retarded individuals. (Grossman, 1977, p. 16)

Some of the more widely used brief screening devices are the Peabody Picture Vocabulary Test, the Slosson Intelligence Test, the Quick Test, and the Columbia Mental Maturity Scale. There is nothing inherently wrong with these devices if they are used very cautiously and for the purpose of screening. In some instances these devices have the advantage of saving time and providing a response format that enables children with physical handicaps or severe expressive language problems to respond to intelligence test items.

The major disadvantages of most of the brief screening devices are poor standardization procedures, limited and/or unrepresentative normative samples, lower reliability, lower validity, limited range of abilities sampled, and limited opportunity to observe qualitative aspects of performance. In most instances the results from such devices are not strictly comparable or interchangeable with the results of the more conventional major instruments, such as the Wechsler Scales, the Stanford Binet, or the McCarthy Scales of Children's Abilities.

The concerns with comparability of scores and appropriate norms are particularly important to *classification* decisions. The data presented in test manuals on the correlation of a brief test with a more conventional test sometimes lead examiners to the erroneous conclusion that the scores are comparable, perhaps even interchangeable. The author of one test mentioned in the preceding paragraph cited correlations greater than .90 between his test and the Stanford-Binet. A correlation of this magnitude indicates that the tests are measuring similar attributes but it does not mean that the scores are comparable. For the test in question, the norms are poor and the scores are based on a ratio IQ scale that has a slightly different mean and a varying standard deviation depending on age. Do these factors create problems in interpreting results?

---

Task 1

*a) How should you interpret an IQ of 65 from a brief screening measure if at that age level the mean for the standardization sample was 98 and the standard deviation was 25? What is the comparable score on the Stanford-Binet? Could this difference influence a classification/ placement decision?*

*b) Do high correlations indicate that scores from two tests are comparable or interchangeable? Inspect the following score distributions carefully and then estimate the correlations between the scores.*

| | Test I<br>*Score Distribution* | Test II<br>*Score Distribution* |
|---|---|---|
| *Subject 1* | 55 | 65 |
| *Subject 2* | 65 | 75 |
| *Subject 3* | 75 | 85 |
| *Subject 4* | 85 | 95 |
| *Subject 5* | 95 | 105 |

____________ *What is your estimate of the correlation?*

____________ *Test 1 standard deviation*

____________ *Test 2 standard deviation*

____________ *Test 1 mean*

____________ *Test 2 mean*

---

Both examples in Task 1 are realistic. There are some brief screening measures that correlate at a very high level with conventional measures of intelligence. However, the scores may not be comparable owing to differences in normative samples, differences in scaled score, or both. To summarize, brief measures of intelligence are useful for some purposes, such as initial screening. They often are misused and misinterpreted. Brief screening measures of intelligence never should be used as the only or primary source of information on intelligence in classification decisions of mild mental retardation.

**Illustration 2. Inappropriate Test.** The selection of an intelligence measure must be based on information about the child. Numerous examples can be cited of a very good test that was used inappropriately. Common sense should prevent most of these abuses. In evaluating these misuses be mindful of two classic assumptions in psychoeducational assessment (Newland, 1980): (a) the child has comparable acculturation to the normative sample, and (b) the child's maximum performance is obtained on ability tests. Moreover, one of the requirements of the PL 94-142 Rules and Regulations is,

> Tests are selected and administered so as best to ensure that when a test is administered to a child with impaired sensory, manual, or speaking skills, the test results accurately reflect the child's aptitude or achievement level or whatever other factors the test purports to measure, rather than reflecting the child's impaired sensory, manual, or speaking skills (except where those skills are the factors which the test purports to measure).

---

Task 2

*The following illustrations are examples of the misuses of a conventional test due to poor test selection. In each case the type of test selection was inappropriate because of the child's characteristics IF a classification/placement decision was being considered. However, each test might have been appropriate IF the purpose was to estimate the effects of a handicap or background characteristics on the child's performance.*

*For each example, specify why the test was inappropriate for a classification/placement decision and suggest more appropriate options.*

*a) Use of the Stanford-Binet with monolingual Spanish-speaking children referred for special education consideration.*

*b) Use of performance or nonverbal tests, such as the WISC-R Performance Scale, with visually impaired children.*

*c) Use of timed tests with children who have neuromuscular handicaps.*

*d) Use of highly verbal measures (e.g., the Stanford-Binet) with hearing-impaired children.*

---

In addition to the obvious examples of misuses provided in the preceding simulation exercise, other examples could be cited. Unfortunately, citing the examples is not merely an academic exercise; instances of at least one, and for some examples, many instances of identical or similar misuses have been observed in practical situations. In each example the selection of the assessment device was inappropriate if an inference about general ability was the purpose of the assessment.

Clearly, in assessments that may lead to a classification of mental retardation, consideration must be given to selecting the very best instruments based on the children's background and abilities to make the required responses. The information needed about the child to make a wise test selection generally is obtained in the early steps of the multifactor assessment.

**Proper Interpretation of Intelligence Test Results**

Intelligence test results often are misinterpreted or overinterpreted when they are used as part of the assessment battery for children referred to special services in schools. Because assessment of intelligence is fundamental in identifying mental retardation, the proper interpretation of intelligence test results is one of the most important tasks of diagnostic personnel.

---

*Task 3*

*Evaluate the following statements on the performance of children on the Wechsler Scales:*

T F 1. *Normal children (i.e., most children) have Verbal and Performance Scale IQ Scores on the Wechsler Scales that are within 9 points.*

T F 2. *Most normal children have fairly flat profiles on the Wechsler Scales; that is, there is relatively little difference among the subtest scaled scores.*

T F 3. *Unique subtest patterns can be identified for mildly retarded persons, and these patterns are useful in making a differential diagnosis.*

T F 4. *Most persons who obtain Full Scale IQ scores in the mildly retarded range have fairly flat profiles across different subtests on the Wechsler Scale.*

T F 5. *A major difference between mildly retarded and learning disabled or emotionally disturbed children is in the amount of variation across the subtest scores on the Wechsler Scales.*

---

The answer to all the items is false. Each statement is based on common misinterpretations of intelligence test data. Results of empirical research and some references for additional study on each statement follow:

1. According to Kaufman's (1976) analysis of the standardization data for the WISC-R, about 50% of children obtain verbal-performance scale differences of 9 points or more. Further, about 25% had verbal-performance discrepancies of 15 points or more. Verbal-performance IQ score discrepancies of this magnitude are reliable differences, but they are not unusual or indicative of abnormal development.

2. Again, the analysis of the WISC-R standardization data by Kaufman (1976) indicates that flat profiles are unusual, and that "scatter" is typical. According to Kaufman, the average amount of variation among the WISC-R subtest scaled scores is 7 points. Therefore, subtest scatter is not a unique characteristic of any diagnostic group.

3. According to Silverstein's (1968) work with the WISC and Reschly and Davis's (1977) with the WISC-R, mentally retarded children do best on Object Assembly and Picture Completion and poorest on Vocabulary, Information, and Arithmetic. However, the differences are rather small, and given the results on normal variation (see No. 2 above), this pattern is not a useful basis or support for a differential diagnosis of mental retardation.

4. Relatively little data are available on this assertion. However, Kaufman (1976) reported no relation between amount of scatter and level of intelligence in a study of test scatter in the standardization sample. Apparently, persons with borderline intelligence or intelligence in the mentally retarded range are no more likely than normal children to have flat profiles.

5. See Nos. 2, 3, and 4 above. If scatter is normal, then it is virtually impossible for scatter or even patterns of scatter to be a unique characteristic of any group. Therefore, diagnoses based primarily on scatter are likely to be inaccurate.

Several excellent sources have been published recently on the proper interpretation of the WISC-R. In addition to the preceding citations, the reader is referred to Kaufman (1979a, 1979b); and Reschly & Reschly, (1979).

**Myths About Intelligence.** Three myths about the meaning of intelligence test results continue to be fairly common, at least among consumers of IQ test results (e.g., parents and teachers).

**Myth No. 1: Intelligence is unitary; IQ tests measure all the important aspects of intelligence.**

Perhaps the most difficult concept to communicate to others is the relation of IQ test results to the construct of intelligence. Part of the difficulty undoubtedly arises from the lack of agreement on the meaning of the construct of intelligence and the simplistic use of the operational definition, "Intelligence is what the IQ tests measure." Clearly, intelligence is more than what the IQ tests measure.

Two of the more prominent theories and definitions of intelligence stress the importance of a variety of competencies, including personality or behavioral factors. Wechsler's global definition of intelligence and Guilford's three dimensional model of intellect are useful devices in our effort to communicate to others that IQ tests do not measure all aspects of intelligence. Analysis of these theories clearly suggests that intelligence is a many-faceted, not unitary, dimension of the individual.

**Myth No. 2: Intelligence or IQ test results are fixed.**

The data on the stability of IQ test results are sometimes misunderstood. IQ test results are fairly stable for groups of individuals after age 6. Moreover, IQ test results are at least as stable for mentally retarded persons, perhaps a bit more stable. However, IQ test results do change for a significant percentage of individuals (at least 20%) by 15 points or more between age 6 and maturity. Considerably larger changes in the magnitude of 30 to 40 points have been reported in rare cases. What do these data mean and how should they influence our interpretations of IQ test results?

Clearly, such data should convince us that IQ, although stable, is not fixed. Large changes, when they occur, usually are associated with significant changes in the individual's environment and/or personal-social adjustment. Large changes in IQ rarely are simple random fluctuations. The fact that IQ test results do change as a function of changes in the individual and/or environment may be seen as evidence for increasing our confidence in them as indications of *current intellectual functioning* (or current academic aptitude). An emphasis on current intellectual functioning probably is the most common interpretation of IQ results.

The federal requirement that a comprehensive reevaluation be conducted at least every three years for students placed in special education programs is an important protection for children in light of the data on stability of measured intelligence. Consumers of IQ tests must be informed that scores (and "ability") can change. Inferences about the future intellectual status of an individual always are tentative.

**Myth No. 3: Intelligence is predetermined by genetic factors; IQ test results reflect genetic potential.**

This myth, that IQ is predetermined by genetic factors is complex. Clearly, the information and problem-solving skills required by IQ test items are learned. However, this fact does not preclude the influence of genetic factors on test scores. Although nearly irrefutable data exist to confirm the influence of genetic factors on measured intelligence, the unanswered (and unanswerable) issues are (a) the amount of influence that is attributable to genetics and (b) the genetic influence on an individual's score.

Mercer (1979a) provided a useful summary of the precise conditions that must be met in order to legitimately interpret differences in the scores of individuals (or groups) as reflections of different levels of innate potential. Individuals (or groups) must have (a) equal exposure to opportunities to learn the information or problem-solving skills measured by the test; (b) equal levels of motivation to learn and reinforcement for learning whatever the test requires; (c) equal familiarity with tests and test-taking situations; (d) equal influence of affective factors (e.g., anxiety, fear, and emotional turmoil) that may interfere with learning or test performance; and (e) equal influence of physical, sensory, or motor abilities that might interfere with test performance or learning. Ensuring that these criteria are met when an individual is given an IQ or other test is virtually impossible.

Considerable research evidence supports the dual influence of heredity and environment on measured intelligence. There is no feasible way to separate these influences when the intelligence test results are interpreted for an individual. The predetermination myth must be kept in mind when test results are communicated to students, parents, teachers, or other persons.

**What Do IQ Tests Mean?** In view of the many cautions that have been expressed thus far, perhaps it is time to consider what IQ tests do measure. If intelligence is not unitary, fixed, or predetermined, is there anything worthwhile that remains?

Ample evidence supports the predictive validity of IQ tests when the criterion is academic performance (Matarazzo, 1972; Sattler, 1974). Results of individual intelligence tests are highly related to results of standardized achievement tests (correlations of .5-.7) and, to a slightly lesser degree, to grades in academic subjects or teacher ratings of academic performance. These relations seem to hold up equally well for white, black, and Hispanic students (Reschly, in press).

The relation of IQ test results to other criteria is far less clear and lower correlations usually are reported. These lower correlations probably are due, in part, to inadequate measures of other criteria, the absence of agreed-upon definitions of "success" in most settings, and the many other factors in addition to IQ that influence competence or success.

IQ test results should be seen as reflecting an important, albeit relatively limited range, of human competencies. Relations with academic achievement certainly are not trivial, and relations with a variety of other variables (e.g., occupational attainment, socio-economic status) are also important (Matarazzo, 1972). (In view of the high correlation of IQ tests with academic performance the term "academic aptitude" is suggested for use in reports as a means of reducing misconceptions and misinterpretations of the test results.) Mercer's (1979) use of the term "School Functioning Level (SFL)," appears to be motivated by a similar concern. The intent behind both efforts is to communicate the fact that IQ test results reflect the likelihood of success in schools. This relationship, although not perfect, is not likely to change unless the goals of schools change radically. As long as schools stress the mastery of the symbols of our culture, such as reading, mathematics, written expression, verbal expression, and so on, our conventional IQ tests will predict school performance reasonably well. IQ test results are not particularly strong indicators of other important aspects of "intelligent behavior," for example, persistence, common sense, and social skills. Changing the name of what present IQ tests measure is not a cure-all for all the abuses and misuses of current tests, but the change in terminology should aid our efforts to ensure the correct understanding of intelligence test results by consumers.

---

*Use Simulation 4 here*

---

**Summary**

The result of properly selected and appropriately administered intelligence tests are useful in predicting academic achievement and in establishing *tentative* expectations for level of academic performance. Care should be taken, however, that no unwarranted assumptions or interpretations are made based on common myths about intelligence or about what intelligence tests measure. When used with other information generated in a multifactor assessment, intelligence test results are valuable in making classification/placement decisions and in identifying children who require a modified educational curriculum. They should never be used as the sole determinant for these decisions.

## Adaptive Behavior

Concern with what is now called "adaptive behavior" is not new. The term "social competence" was used prior to about 1960 to refer to approximately the same construct. Social competence or adaptive behavior also has been a fundamental concept in efforts to describe and explain the phenomenon of mental retardation. Further, Alfred Binet's discussions of the construct of intelligence included a theme that we might now refer to as adaptive behavior.

Although the construct is not new, a number of recent events have added emphasis to the use of adaptive behavior data in special education classification and placement decisions. The revisions of the AAMD *Manual on Terminology and Classification* in 1961 and 1973 (i.e., Grossman, 1973; Heber, 1961) reflected this increasing emphasis. The "normalization" effort to integrate institutionalized mentally retarded persons into community settings is a second major influence stressing adaptive behavior. From this perspective, adaptive behaviors are viewed as the "reversible" features of the more severe levels of mental retardation (Leland, 1978). A somewhat unrelated trend is the emphasis on nonbiased assessment that resulted from litigation and legislation in the 1970s. Adaptive behavior from this perspective is seen as a means to reduce the emphasis on intelligence test results, to provide more equitable assessment for minorities, and to alleviate the overrepresentation of minorities in special education programs for the mildly retarded (Coulter & Morrow, 1978).

In view of the diverse influences and different purposes underlying the recent upsurge of interest in adaptive behavior, it is not surprising that much confusion exists about the measurement and use of adaptive behavior data. Adding to this confusion is the implication in recent federal legislation that adaptive behavior data must be considered in *all* special education placement decisions. Perhaps the best recent source of information on adaptive behavior is the book by Coulter and Morrow (1978), cited in the Recommended Readings for this section, which has an excellent discussion of available measures, possible uses, and unresolved issues.

### Adaptive Behavior and Definitions of Mental Retardation

For approximately two decades the AAMD definition of mental retardation has included the dimensions of intelligence and adaptive behavior. However, the emphasis on adaptive behavior was increased in the 1973 version. The 1961 version described mental retardation as subaverage general intellectual functioning which is associated with impairment in adaptive behavior. The 1973 and 1977 versions placed more emphasis on adaptive behavior by changing "associated" to "*existing concurrently*." This change—placing relatively equal emphasis on both dimensions of mental retardation—along with the subtle changes in the conception of adaptive behavior (cf. the 1961 and 1973 versions) constitute difficult challenges for diagnostic personnel.

Most educational definitions of mental retardation tend to include both the intelligence and adaptive behavior dimensions. According to a 1976 survey, about half the states required assessment of adaptive behavior for one or more special education classifications, usually mental retardation (Coulter & Morrow, 1978). A number of additional states reported efforts to add adaptive behavior to the state definition of mental retardation. However, the majority of states did not define adaptive behavior; much confusion was reported about definition, domains of adaptive behavior, and availability of measures. Although the status of adaptive behavior in special education undoubtedly varies from state to state, the trend is toward more emphasis on this dimension, at least for mentally retarded students.

### Conceptions of Adaptive Behavior

Adaptive behavior is a complex construct that shares some similarities with the construct of intelligence. Both have a diversity of theories, measurement operations, and practical applications. A major difference between the two is the relatively little research on adaptive behavior.

One of the most influential definitions and descriptions of adaptive behavior is provided in the AAMD *Manual*. The AAMD conception and criteria for adaptive behavior during the school-age years changed in subtle ways from 1961 to 1973. Consider the following description from the 1961 revision:

> Adaptive behavior refers primarily to the effectiveness of the individual in adapting to the natural and social demand of his environment. Impaired adaptive behavior may be reflected in: 1) maturation, 2) learning, and/or 3) social adjustment. These three aspects of adaptation are of different importance as qualifying conditions of mental retardation for different age groups.
>
> Learning ability refers to the facility with which knowledge is acquired as a function of experience. Learning difficulties are usually most manifest in the academic situation and if mild in degree may not even become apparent until the child enters school. Impaired learning ability is, therefore, particularly important as a qualifying condition of mental retardation during the school years. (Heber, 1961, pp. 3-4)

On the basis of this 1961 description, school-age children's adaptive behavior was determined by their academic competence; thus, a diagnosis of mental retardation for such children could be based only on intelligence, classroom academic performance, and results of standardized achievement tests. Other characteristics and behaviors specified in current conceptions of multifactor assessments should have been used and, fortunately, they often were considered in the classification/placement decisions for children with mild mental retardation. Nevertheless, the clear implication of the 1961 revision was that academic performance was the most important index of school-age children's adaptive behavior. With considerable justification one could argue that up to 1973 when the AAMD *Manual* was revised, diagnostic personnel in the schools were assessing adaptive behavior as it was conceptualized at that time.

The changes in conception of adaptive behavior for school age children in the 1973 and 1977 revisions of the AAMD *Manual* which are vitually identical, are illustrated in the following quotes:

> *Adaptive behavior* is defined as the effectiveness or degree with which an individual meets the standards of personal independence and social responsibility expected for age and cultural group (Grossman, 1977, p. 11).
>
> Since these expectations of adaptive behavior vary for different age groups, *deficits in adaptive behavior* will vary at different ages. These may be reflected in the following areas:
>
> During *infancy and early childhood* in:
>
> 1. Sensory-motor skills development
> 2. Communication skills (including speech and language)
> 3. Self help skills
> 4. Socialization (development of ability to interact with others)
>
> During *childhood and early adolescence* in:
>
> 5. Application of basic academic skills in daily life activities
> 6. Application of appropriate reasoning and judgment in mastery of the environment
> 7. Social skills (participation in group activities and interpersonal relationships)
>
> and
>
> During *late adolescence and adult life* in:
>
> 8. Vocational and social responsibilities and performance
>
> During infancy and early childhood, sensory-motor, communication, self help, and socialization skills ordinarily develop in a sequential pattern reflective of the maturation process. Delays in the acquisition of these skills represent potential deficiencies in adaptive behavior and become the criteria for mental retardation.
>
> The skills required for adaptation during childhood and early adolescence involve complex learning processes. This involves the process by which knowledge is acquired and retained as functions of the experiences of the individual. Difficulties in learning are usually manifested in the academic situation but in evaluation of adaptive behavior, attention should focus not only on the basic academic skills and their use, but also on skills essential to cope with the environment, including concepts of time and money, self-directed behaviors, social responsiveness, and interactive skills.
>
> In the adult years, vocational performance and social responsiblities assume prime importance as qualifying conditions of mental retardation. These are assessed in terms of the degree to which the individual is able to maintain himself independently in the community and in gainful employment as well as by his ability to meet and conform to standards set up by the community. (Grossman, 1977, pp. 13-14)

The recent revisions of the AAMD *Manual* placed more emphasis on and broadened the concept of adaptive behavior during the school-age years. Contrary to some recent trends in conceptions and measures of adaptive behavior, the AAMD conception *does* continue to include performance in academic settings as an important component of adaptive behavior during the school-age years. For children in this age group school performance is a necessary part of the construct of adaptive behavior (see following subsection). However, performance in other social settings also must be considered.

A thorough review of conceptions of adaptive behavior (Coulter & Morrow, 1978) reveals major differences in relation to their *purposes* or intended use and to the domains of behavior included within the constructs.

### Analysis of Conceptions of Adaptive Behavior

Among the conceptions of adaptive behavior, a number of common features and differences can be identified. They have important implications for the collection and use of adaptive behavior data by diagnostic personnel. Two features—(1) emphasis on age-appropriate criteria and (2) recognition of the importance of the cultural context—are common to most definitions of adaptive behavior. Conceptions of adaptive behavior differ in their answers to three other considerations—(3) whether adaptive behavior competencies are situational or generalized, (4) what domains of competencies are included, and (5) what is the purpose of the assessment of adaptive behavior.

1. **Developmental.** Every major definition or description of adaptive behavior specifies or implies different standards, depending on the age of the individual.
2. **Cultural Context.** Most definitions of adaptive behavior include the recognition of cultural influences. Competene is judged within a cultural context. Different expectations exist and varying competencies are expected in different cultures.

The other dimensions on which conceptions of adaptive behavior may differ could be anticipated from careful analysis of the implications of these first two features.

3. **Situational or Generalized.** Are adaptive behavior competencies global or situation specific? Most conceptions of adaptive behavior at least imply a situational perspective rather than a generalized trait view. The situational view holds that at a particular age a person may be competent in some domains while he/she is deficient in others. Both the generalized trait and situational views imply that a person can be incompetent in adaptive behavior at one age but competent at another (assuming that the criteria change).
4. **Domains.** The domains in adaptive behaviors reflect some commonalities as well as differences. Nearly every conception includes the notion of self-maintenance or independent functioning (e.g., meeting one's physical, safety, and health needs). Most conceptions include the dimension of interpersonal relationships (e.g., getting along with peers and adults), and some conceptions also include intrapersonal competencies (e.g., freedom from debilitating emotional turmoil). A third common domain is social responsibility (e.g., meeting the expectations of others within particular cultural contexts). Activities or social roles that may be included in this domain are vocational/career competencies, constructive participation in family, neighborhood, and community activities, and assumption of responsibility for the support and care of others.

   A fourth domain, one on which conceptions differ, stresses the importance of cognitive competencies and communication skills. As we have seen, some conceptions (and instruments; see following subsection) place considerable stress on language competence, cognitive development, and academic achievement. Whether this domain is included in our conception of adaptive behavior has significant implications for the classification/placement decisions for mildly retarded students.
5. **Purpose.** Conceptions of and instruments used to measure adaptive behavior differ according to their function in the assessment/intervention process. The two major purposes for measuring adaptive behavior are classification/placement and program planning/intervention.

   The diagnostic questions associated with these two purposes are quite different. The question for classification/placement decisions is whether the person's adaptive behaviors are sufficiently discrepant from normal to justify classification as retarded. The question in programming/intervention decisions is whether certain behaviors prevent a person from being placed in a community rather than an institutional setting. The instruments developed for these two purposes are quite different.

**Assessment of Adaptive Behavior**

The purpose of assessment (i.e., the decision that must be made about or with a student) is the basis for the selection and use of formal or informal data collection procedures (Salvia & Ysseldyke, 1978). Clarifying the purpose through explicit statements of the decisions to be made is particularly important in the assessment of adaptive behavior.

If the purpose of assessment is program planning/intervention for retarded students, the currently available adaptive behavior instruments are reasonably adequate for most ages. Some instruments have been developed carefully with rigorous measurement and statistical criteria applied to the selection of items. A sample list of some of the more widely used instruments, reprinted from Oakland and Goldwater (1979), is given in Table 3.

**Table 3**

**Some Instruments for the Measurement of Adaptive Behavior in School Children**

| | | | | | | | | | | | | | Population type (and size) | | Purpose | | | Examiner | | | Respondent | | | Reliability & validity Data Available | SCORES | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Phys. Deve./Sensory Motor/Locomotion | Self-Direction | Language and Communication | Vocation and Occupation | Economic | Social | Self-Help/Independent Functioning/ Self-Maintenance | With Peers | In School | In the Family | In the Community | Age Range | Clinical | School | Screening | Placement | Programming | Teacher | Diagnostician | Paraprofessional | Child | Teacher | Parent/Family | Yes/No | G.E./Age | File | Scales Score | Administration Time |
| AAMD Climical (Nihira, et.al., 1974 | X | X | X | X | X | X | X | | | | | 3-Adult | X | | | X | X | | X | X | | | X | Yes | | X | | 1 hr. |
| AAMD School Lambert, et. al., 1974 | X | X | X | X | X | X | X | | | | | 7 - 13 | | X | | X | X | X | X | X | | X | X | Yes | | X | | 1 hr. |
| Cain Levine (Cain, et.al., 1963) | | X | X | | | X | X | | | | | 5 - 13 | | | | X | | X | X | X | | | X | Yes | | X | | 20' |
| California Preschool (Levine, et. al., 1969) | | | X | | | X | | | | | | 2 - 5 | | | X | | | X | X | X | | X | X | Yes | | X | | 20' |
| Camelot (Foster, 1974) | X | X | X | X | X | X | X | | | | | 2 - Adult | | | | X | X | X | X | X* | | X | X | Yes | | | | 1 hr. |
| Adaptive Behavior Inventory for Children (Mercer & Lewis, 1978) | | | | X | X | X | X | X | X | X | X | 5 - 11 | | X | | X | | | X | X* | | | X | Yes | | X | X | 1 hr. |
| Preschool Attainment Record (Doll 1966) | X | X | X | | | X | X | | | | | birth-7 | | | | X | | X | X | X | | X | X | No | X | | X | 20' |
| Vineland (Doll, 1965) | X | X | X | X | X | X | X | | | | | birth-25 | X | | | X | | | X | | • | X | X | Yes | | | | 20' |

Source: Reprinted from T. Oakland & D. Goldwater Asseement and interventions for mildly retarded and learning disabled children. In G. Phye & D. Reschly (Ed.) SCHOOL PSYCHOLOGY: PERSPECTIVES AND ISSUES. New York: Academic Press 1979

*with extensive training in interviewing

School psychologists selecting an adaptive behavior measure are usually concerned with choosing an instrument that will contribute data to a nonbiased assessment of a school-age child. Of the adaptive behavior measures listed in Table 3, only two are designed specifically for normal, borderline, and mildly retarded school-age populations: the AAMD School Version and the Adaptive Behavior Inventory for Children (ABIC). Although adaptive behavior is one of the key areas in the multifactor assessment scheme developed by Tucker (1977) and mentioned in the PL 94-142 Rules and Regulations, the present level of technology in measuring adaptive behavior for classification decisions is characterized well in the following quotes:

> The inclusion of adaptive behavior in nonbiased assessment by the use of tests or scales to facilitate comparison of a child with his/her peers is not yet perfected. (CORRC, p. 20, Undated report distributed in 1979)
>
> Presently, the assessment of adaptive behavior through clinical interviews and observations of the child's behavior in other social systems represents the major alternative for pupil appraisal professionals, if the goal of assessment is primarily placement. Until psychometric technology provides a variety of suitable and more objective behavior measures, the more informal, and thereby subjective, methods will remain in wide use. (CORRC, p. 21, Undated report distributed in 1979)

Problems in assessing adaptive behavior also received prominent mention in the AAMD *Manual* (Grossman, 1979, pp. 20-21): (a) the frequent discrepancies between levels of adaptive behavior and intelligence of mildly retarded persons; (b) the unavailability of adaptive behavior instruments sufficiently precise to establish a definite cut-off score (e.g., minus two standard deviations from the mean); and (c) the major limitations of most available instruments (e.g., inadequate norms and item content). In view of these and other limitations, Grossman suggested that the assessment of adaptive behavior include clinical judgment.

Inasmuch as the available technology for the assessment of adaptive behavior of normal and mildly retarded children leaves much to be desired, a considerable amount of additional work on instrument development and research is clearly needed.

The picture suggested in the preceding quotations from CORRC may be a bit too negative, however. Instrument development and research in recent years can be applied to the assessment of adaptive behavior in classification/placement decisions. The judicious use of the results of these instruments along with informal sources of data on adaptive behavior should become a part of the comprehensive evaluation conducted prior to classification/placement decisions.

Because the adaptive behavior instruments that are useful for assessment in classification/placement decisions for mildly handicapped students are relatively new, their selection and evaluation are especially challenging. Simulation 5 is designed to provide information to help guide your choices.

---

*Use Simulation 5 here.*

---

### Review of Adaptive Behavior Measures for Mildly Retarded Students

**AAMD Adaptive Behavior Scale - Public School (ABS-PS).** The most important influences leading to the development of the ABS-PS were legal requirements in California on the classification/placement of students in EMR programs. Other purposes, such as providing information for educational programs and remediation, also were cited by the authors (Lambert, Windmiller, Cole, & Figueroa, 1975).

The items on the ABS-PS are a subset of items from the AAMD Adaptive Behavior Scale - Clinical (ABS-C). The ABS-C was developed from extensive studies of deficit behaviors among institutionalized mentally retarded persons. The purpose of the ABS-C was to pinpoint behaviors that prevent the placement of severely retarded persons in community settings. When these behaviors are identified, the focus, then, is on remediation and, eventually, placement in less restrictive settings. The critical point is that the items on the ABS-C were selected from studies which were undertaken to improve program planning/intervention of severely retarded institutionalized persons. The content of the Public School version is drawn from and, therefore, is similar to the Clinical version. It appears to be more appropriate for lower functioning children in the EMR range.

The ABS - Public School is divided into two major sections. Section I might be termed "adaptive behaviors" inasmuch as higher scores here indicate higher social functioning. Section II might be called "maladaptive behaviors" because higher scores indicate lower social functioning. The nine domains with 56 items on Section I include Independent Functioning, Physical Development, Economic Activity, Language Development, Numbers and Time, Vocational Activity, Self-Direction, Responsibility, and Socialization. A sample item from the Shopping Skills area of the Economic Activity Domain follows:

30. Errands (Circle only one)

| | |
|---|---|
| Goes to several shops and specifies different items | 4 |
| Goes to one shop and specifies one item | 3 |
| Goes on errands for simple purchasing without a note | 2 |
| Goes on errands for simple purchasing with a note | 1 |
| Cannot be sent on errands | 0 |

Section II comprises 39 maladaptive behavior items, including the following 12 domains: Violent and Destructive Behavior, Anti-social Behavior, Rebellious Behavior, Untrustworthy Behavior, Withdrawal, Stereotyped Behavior and Odd Mannerisms, Inappropriate Interpersonal Manners, Unacceptable Vocal Habits, Unacceptable or Eccentric Habits, Hyperactive Tendencies, Psychological Disturbances, and Use of Medications. Item 32 of the Hyperactive Tendencies Domain is as follows:

32. Has Hyperactive Tendencies

| | Occasionally | Frequently |
|---|---|---|
| Talks excessively | 1 | 2 |
| Will not sit still for any length of time | 1 | 2 |
| Constantly runs or jumps around the room or hall | 1 | 2 |
| Moves or fidgets constantly | 1 | 2 |
| Other (Specify ____________) | 1 | 2 |
| _____ None of the above | | |
| | Total____________ | |

The recommended respondent for the ABS-PS is the child's classroom teacher. Respondents are allowed to infer or, if necessary, to guess the child's competencies, particularly those that are displayed outside of school.

The norms for the ABS-PS are based on a California sample of 2600 school-age children between the ages of 7 and 13. Separate norms are provided by class placement (regular vs. types of special classes) for both sections. In addition, separate norms by ethnicity and sex are provided for Section II.

The interpretation of the ABS-PS is based on the comparison of the individual's profile of percentile ranks to the modal profiles of children who are placed in different educational programs. No standard scores are provided for the domain scores, and no over-all score for the major sections is available.

The ABS-PS may be a useful adjunct to clinical judgment in classification/placement decisions and to a lesser degree in program planning/intervention decisions. Major weaknesses in the instrument exist, however. The content validity of the items is questionable in view of the original purpose of the ABS-Clinical version. The item format requires a considerable degree of inference or guessing. The respondent is the teacher who usually has little information about social role performance outside of school. Finally, the method of interpretation—comparing profiles—is highly subjective in many cases.

**The Adaptive Behavior Inventory for Children (ABIC)** was developed with the explicit purpose of improving classification/placement decisions for mildly retarded children (Mercer, 1979a). The ABIC reflects a strong social system perspective with the emphasis on how children function in different settings and different social roles.

The ABIC items were selected on the basis of intensive interviews with mothers of children between the ages of 5 and 11. Most of the 242 items are age graded. A basal and ceiling procedure is used to administer the ABIC. The domains covered by the measure are Family, Community, Peer Group, Nonacademic School, Earner/Consumer, and Self-Maintenance.

Sample items from each ABIC domain are as follows:

| **Domain** | **Item** |
|---|---|
| Family | 147. When ____ cannot have what he/she wants immediately, how often does he/she get angry and fuss about it?<br>0 most of the time<br>1 sometimes, or<br>2 almost never |
| Community | 142. When visiting relatives or friends outside the neighborhood, does ____ usually<br>0 go with an older person<br>1 go with children his/her age, or<br>2 go alone? |
| Peer Relations | 144. How often does ____ meet and play with his/her friends at a special place like a vacant lot, a park, the street, the school bus stop, or a courtyard?<br>1 sometimes<br>0 seldom or never, or<br>2 often |
| Nonacademic/ School | 132. How often does ____ take his/her school supplies and books to school without being reminded?<br>1 occasionally<br>0 seldom, or<br>2 regularly |
| Earner/ Consumer | 140. Does ____ make correct change for a dollar?<br>2 without help<br>1 only with help, or<br>0 not at all? |
| Self-Maintenance | 143. Does ____ order food at a restaurant<br>2 without help<br>1 with some help, or<br>0 does someone order for him/her? |

The ABIC is administered as a structured interview. The primary caretaker of the child, typically the mother, is the preferred respondent. The mother chooses among three possible responses for each item.

Standard scores with a mean of 50 and standard deviation of 15 are provided for each domain. The average of these standard scores is used as a composite or global index of adaptive behavior. In addition, three other scores are provided: The Veracity Scale attempts to detect a respondent's tendency to fake a good response set. The "No Opportunity" and "Not Allowed" responses indicate the amount of restriction that is placed on a child. Finally, the "Don't Know" responses are regarded as an indication of the amount of knowledge the respondent has about the child's activities. If critical values are exceeded on the three ancillary scales, interpretation of the other scores is not recommended.

The ABIC norms are based on a stratified random sample of 2100 California school children between the ages of 5 and 11. Stratification variables included sociocultural group (Anglo, black, and Hispanic), size of community, and gender.

The ABIC is one of the few instruments in which the entire design (from item selection to standardization) was directed toward classification/placement decisions for normal, borderline, and mildly retarded children. The face validity of the items appears to be good. The type of derived scores appear to be appropriate for classification/placement decisions. The ancillary measures safeguard against interpreting invalid information. The primary type of information derived from the ABIC is related to children's social role performance *outside of school from the perspectives of parents* (or primary caretakers).

Although the ABIC is one of the best instruments for assessing the out-of-school adaptive behaviors of normal or mildly handicapped children, a number of weaknesses should be recognized: The age range is limited to 5-11 years. The norms are based entirely on California school-age children. The accuracy of these norms in other settings still needs to be established. An important domain of school-age children's adaptive behavior—academic role performance—is not included on the scale and is de-emphasized in Mercer's conception of adaptive behavior. Finally, practical considerations of time and resources may limit the use of this method of assessing adaptive behavior.

**The Vineland Social Maturity Scale (VSMS)** (Doll, 1953) is one of the oldest measures of social competence (adaptive behavior) and it continues to be used widely (Coulter & Morrow, 1978). One reason for the current use of the VSMS is that other scales are limited in age range or were not available until very recently.

The VSMS is a loosely structured interview that requires considerable administrative skill. It attempts to measure social competence of persons from birth to age 30 years. Thus the items vary considerably in sophistication and ease of administration. The domains of behavior covered by the measure are Self-Help General, Self-Help Eating, Self-Help Dressing, Locomotion, Occupation, Communication, Self-Direction, and Socialization. The VSMS yields a composite socre which can be transformed into a Social Quotient (SQ): the ratio of Social Age divided by chronological age and then multiplied by 100. The standard deviation of the SQ varies considerably from age to age. The norms for the VSMS are based on rather restricted samples of individuals who were assessed in 1935.

The VSMS is in rather desperate need of revision and "renorming," an activity that currently is underway and which may substantially improve the scale. For older students, it provides some information that can be used to supplement clinical judgment of adaptive behavior. Direct use of SQ scores in classification/placement decisions is probably inappropriate for a variety of reasons (poor norms, limited sample of behavior, etc.).

**The Children's Adaptive Behavior Scale (CABS)** is a recently developed adaptive behavior scale which reflects some innovative approaches. The CABS (Richmond & Kicklighter, 1980) is administered directly to the child rather than to a third party respondent. The items on the CABS are organized around the typical domains of Language Development, Independent Functioning, Family Role Performance, Economic-Vocational Activity, and Socialization.

**In contrast to other adaptive behavior measures, the CABS emphasizes the cognitive competencies which are required for various adaptive behaviors. For example, on the Independent Functioning Domain one of the items is "Where could you find a doctor?" In the Socialization Domain one of the items is "What should you say if someone gives you a piece of candy?" The norms for the CABS are based on rather restricted samples of slow learning and EMR students.**

**Relatively little is known about the CABS. It is likely that considerable research will be conducted with this scale in the future. For the time being the CABS should be used cautiously, if at all, pending research on its psychometric characteristics.**

---

*Use Simulation 6 here.*

---

## Nonbiased Classification/Placement in Mild Mental Retardation

**Before considering ways to combine information on intelligence and adaptive behavior in classification/placement decisions, a more accurate conception of the problem and a clear perspective on the entire issue of nonbiased assessment is needed.** ***The perspective suggested here does not have the official approval of any agency.*** **It is offered as one approach to deal with the challenge of providing fair and useful assessment and placement services to all children.**

### Analysis of Data on Placement Bias

**That minority students were more likely to be classified and placed in special class programs for the mildly retarded was and, perhaps, is apparent from the enrollment data of school districts throughout the United States. Data on the enrollment of minorities in other special education categories has not been analyzed so frequently, but it appears that minorities are not overrepresented in the low-incidence or more severe handicapping conditions. However, minorities may be overrepresented in other programs for children with mildly handicapping characteristics (e.g., emotional or learning disabilities). Still, mild mental retardation has been the principal area of concern regarding overrepresentation**

**A number of court cases have dealt with the issue of overrepresentation (*Diana*, 1970; *Guadalupe*, 1972; *Larry P.*, 1972; *Mattie T.*, 1979). Most have been resolved through consent decrees that specified actions designed to eliminate the overrepresentation. The one exception is the *Larry P. vs. Riles* case; it resulted in a full trial after the court ordered injunctions in 1972 and 1974. The trial, which was very complex, extended from about October 1977 to May 1978. In October 1979 Judge Peckham's *Opinion* ordered the California State Department of Education to (a) eliminate the use of intelligence tests with black children for the purpose of mental retardation classification/placement; (b) reevaluate all black children currently classified as mentally retarded, and (c) eliminate overrepresentation of minorities in special class programs. (For additional information on the *Larry P.* case see Bersoff's section of this module and vol. 9, issue no. 2 of the *School Psychology Review*).**

---

### *Task 4*

### *Analysis of Data Indicating Placement Bias*

1. *The following data were undisputed facts in the Larry P. vs. Riles case. The enrollment in the San Francisco Public Schools in the early 1970s was 28% black, but enrollment in special EMR classes was 66% black. A similar pattern of overrepresentation was apparent in data for the entire state of California where 10% of the total population was black and 25% of the enrollment in special EMR classes was black.*

   *Based on these data, estimate the percentage of black children placed in special EMR classes.*
   - *a. 1 to 3%*
   - *b. 3 to 5%*
   - *c. 5 to 10%*
   - *d. 10 to 33%*
   - *e. over 33%*

   *Analysis of California enrollment data*

| | **1968-69** | **1976-77** |
|---|---|---|
| *Total student enrollment* | *4,500,000* | *4,380,000* |
| *Total black enrollment (10%)* | *450,000* | *438,000* |
| *Total enrollment in special EMR classes* | *57,148* | *19,289* |
| *Black enrollment in special EMR classes* | *14,573*<br>*(25.5%* | *4,899*<br>*(25.4%)* |

   *What percent of total student enrollment was in special EMR classes?*

   *1968-69: 57,148 — 4,500,000 = 1.3%*
   *1976-77: 19,289 — 4,380,000 = 0.4%*

   *What percent of black children were in special EMR classes?*

   *1968-69: 14,573 — 450,000 3= 3.2%*
   *1976-77: 4,899 — 438,000 = 1.1%*

2. *Analyze enrollment data for your district.*
   - *a) What percent of the student population is minority? What percent of the special education population is minority in the separate categories for mental retardation, learning disabilities, and emotional disabilities?*
   - *c) What percent of the minority population is in each of the above three categories?*
   - *d) What percent of the minority population is in different types of special education service options (e.g., special class vs. resource)?*
   - *e) Given an overrepresentation, is there a "compelling educational justification" for the overrepresentation? What is the justification?*

---

**This exercise on overrepresentation makes the point that enrollment data often are misunderstood. Clearly, even in situations in which overrepresentation appears to be quite high, the *actual* percentage of minority students placed in programs for the mildly retarded may be fairly low. Overrepresentation is a legitimate problem, particularly if the special education programs are perceived to be poorly conducted, ineffective, or provided under conditions that unnecessarily promote the segregation of student groups. However, the re-analysis of the overrepresentation data given in Task 4 leads to a more accurate conception of the problem, which in turn, may assist us to develop more reasonable courses of action.**

**The litigation on alleged placement bias had a significant influence on state and federal legislation (Turnbull, 1978). The following quotes from PL 94-142 are characteristic of the strongly worded statements that have appeared in both litigation and legislation.**

> **Testing and evaluation materials and procedures used for the purposes of evaluation and placement of handicapped children must be selected and administered so as not to be racially or culturally discriminatory.**
>
> **Tests and other evaluation materials include those tailored to assess specific areas of educational need and not merely those which are designed to provide a single general intelligence quotient.**
>
> **No single procedure is used as the sole criterion for**

> determining an appropriate educational program for a child. *(Rules and Regulations, Federal Register, August 23, 1977).*

An assumption in current legislative trends is that intelligence tests are primarily responsible for placement bias and other information on children is either not developed or ignored. Although the general mandate to reduce bias in assessment is clear and seemingly unequivocal, there is a startling absence of criteria or guidelines in the litigation and legislation. The fact that a clear-cut definition of bias and operational guidelines have not been formulated is not surprising when we consider the contents of the following section.

**Bias in Intelligence Tests**

One of the most troublesome issues in school psychology and special education is the allegation that conventional IQ tests are biased against children from certain racial-ethnic groups. We have a clear mandate in PL 94-142 to eliminate discrimination in assessment procedures and placement decisions. Several extensive reviews of definitions of bias and reviews of research on specific tests or procedures have appeared recently (see the chapters by Jones, Ysseldyke, Mercer, and Page in Morra, 1978; Oakland, 1977; Reschly, 1979; in press). Bias in assessment is defined variously. Conclusions on the bias of a specific instrument, assessment procedure, or placement process will vary depending on one's definition. Just as a test's validity is neither all or none (i.e., is conditional), bias in tests also is conditional. Judgments of bias depend on such features as the definition of bias used, the characteristics of the child being tested, and the purpose of the assessment. There is no universally agreed-upon definition of bias and such agreement is highly unlikely in the foreseeable future.

---

*Use Simulation 7 here.*

---

The simulation you just completed was designed to illustrate the different definitions of test bias and the various conclusions following the different definitions. Some results may be surprising. With few exceptions, conventional standardized tests, such as intelligence tests, appear to measure the same attribute in about the same way, with comparable reliability and validity, regardless of the group membership of the persons being tested. Moreover, attempts to solve the bias problem by changing the tests or developing culture-free or culture-fair tests have not been successful (Anastasi, 1976). The most reasonable solution to the problem of bias in assessment is found in *how* we use the test results, an issue that is discussed in the last part of this section.

**Perspective on Nonbiased Assessment**

Currently, the use of assessment procedures and classification/placement decisions for minority students is controversial and emotionally charged. We have seen in this section that we have no clear-cut solution or explicit criteria to guide our efforts. The present highly ambiguous situation is likely to continue in the foreseeable future. Perhaps one of the most important characteristics of special educators and school psychologists is the ability to tolerate ambiguity and, even more important, is the capability of taking decisive and constructive action on behalf of children despite the ambiguities.

A better perspective on nonbiased assessment can be developed by considering the entire purpose of our professional roles. Simple and narrow-minded avoidance of "placement bias" in special education is of dubious value to the minority students who are failing in regular classrooms. The broader issue is providing more effective educational interventions for minority children and youth. Stubborn defense of the current pattern of overrepresentation or strict avoidance of so-called "placement bias" are both "non-solutions" to the challenge of nonbiased assessment. Against this background, the following definition of bias in assessment was developed:

> Assessment which does not result in effective interventions should be regarded as useless, and biased or unfair as well if ethnic or racial minorities are differentially exposed to ineffective programs as a result of assessment activities. (Reschly, 1979)

The two essential components of this definition of test bias are *usefulness* and *fairness*. The first—usefulness—is seen in assessments resulting in effective interventions that improve skills and competencies and, thereby, enhance opportunities. This goal should continue to be paramount for school psychologists and special education services. The usefulness of all assessment procedures (including intelligence tests) should be determined on the basis of their contributions to this goal. It is acknowledged, in some instances, that assessments lead to accurate diagnoses for which no known effective interventions can be found. These diagnoses still may be valid (in the sense of validity used by Cromwell, Blashfield, & Strauss, 1975) if they improve the estimations of the prognosis or contribute to the prevention of the condition in future cases. However, if effective interventions cannot be developed, the accurate prognostic estimates or the prevention of the condition in the future rarely are of benefit to the individual being assessed.

Fairness is related closely to the notion of usefulness. Assessments and accompanying diagnoses are interpreted as biased or unfair if they result in the overrepresentation of minorities in programs that are *ineffective*, or if no interventions at all are planned. Under such circumstances, the diagnosis may be accurate and the assessment conducted competently, but there is no benefit to the individual. Moreover, if a negative connotation or stigma is associated with a particular diagnosis, and if this diagnosis occurs more often with minority children, then the assessments leading to that diagnosis are regarded as biased or unfair. On the other hand, assessments that lead to the accurate description of current behaviors, to diagnoses that are essentially summary statements of these behaviors, and to effective interventions should be regarded as fair or unbiased *regardless* of the racial-ethnic composition of the student group. Over- or under-representation of minorities in various classifications or programs is not sufficient to establish bias in this conception.

A major problem with the approach suggested here is the possible negative effect of formal special education labels. During the past decade much has been written about labeling effects. This literature indicates that (a) people who bear the labels often find them aversive and unacceptable, and (b) many people (including some professional personnel) misinterpret the meaning of the labels. Other alleged effects of labels *on persons* are largely speculative, with little or no empirical foundation (MacMillan, Jones & Aloia, 1974).

The need for considerable caution by diagnostic personnel in making classification/placement decisions is clear. Other educational remedies must be attempted prior to formal labeling, and formal labeling must be based on extensive collection of data with due process procedures observed carefully. However, simply to avoid labeling is naive. Most children referred to special education diagnostic teams already have been labeled informally in classrooms (e.g., slow learners or "dumb kids"); many children who are referred have serious learning problems, some of a long-standing nature; and other remedies in the context of regular education have been (or should have been) attempted. In the typical situation, the issues are far more complex than the simple avoidance of formal labels.

The ultimate criteria are whether the formal label accurately describes the current status of a child and whether the services that are provided as the result of the classification process benefit the child. Risks must be weighted against possible benefits. Do the special services provided justify the risks associated with the label?

A number of factors that can be identified as prerequisites to achieving fairness in assessment use the effectiveness criterion (Reschly, 1979). Procedural safeguards, such as due process and informed consent, clearly are important (see Bersoff's section of this module). The contributions of all participants in multidisciplinary teams, which represent diverse perspectives on human development and behavior, are necessary to enhance the likelihood of developing comprehensive perspectives on problems and their possible solutions. Most important to nonbiased assessment is the delivery of effective programs in the least restrictive environment. If we can meet these criteria, then we minimize the possibility of litigation over the disproportionate placement of minorities in special education programs. If we are unable to meet them, we will (and should) have ample reason to be frightened about litigation.

**Combining Intelligence and Adaptive Behavior Data**

In addition to the other data from the multifactor assessment, the information on adaptive behavior and intelligence is particularly important in classification/placement decisions for mildly retarded students. How adaptive behavior is conceptualized and measured, along with the available special education service options, significantly influences classification/placement decisions.

The adaptive behavior dimension for school-age children should be conceptualized as two separate components: (a) performance in the school setting, with primary emphasis on academic achievement in the classroom, and (b) role performance in social systems, such as home, neighborhood, and community, outside the school. Separation of the adaptive behavior dimension into two components is necessary because recently published data suggest that adaptive behavior in academic settings and social role performance outside of school may be largely unrelated (Mercer, 1979b).

The inclusion of social role performance in public school in our conception of adaptive behavior is consistent with the description of adaptive behavior for school-age children in the AAMD *Manual*. Two nearly universal features of conceptions of adaptive behavior are age-appropriate criteria and cultural context Developmental task theory stresses the importance of academic performance for children and youth between the ages of 5 to 16 or 18 in our culture. Academic role performance is an important cultural expectation common to all major groups. If adaptive behavior is "the way an individual performs those tasks expected of someone his/her age in his/her culture," then academic performance must be included in any comprehensive view of the construct of adaptive behavior.

Social roles and social systems outside of school also are important domains of development. Again, the conception of adaptive behavior in the AAMD *Manual* and developmental task theory form the foundation for this second component of adaptive behavior. During the school-age years, children perform a variety of social roles of increasing complexity in various social systems. To ignore a child's strengths and weaknesses in social systems outside the school constitutes a serious deficiency in our view of adaptive behavior.

Classification and placement decisions for mildly retarded students should be based on information from both components of adaptive behavior as well as from the dimension of intelligence. A rationale for considering the two dimensions of adaptive behavior is summarized in Table 4, and a flow chart for assessing and combining the adaptive behavior information with intelligence data is shown in Table 5.

The different combinations of adaptive behavior and intelligence hold implications for classification and placement decisions. Adaptive behavior-school (AB-S) should be based on complete educational evaluations, including observation in the classroom, examination of samples of daily work, interviews with the teacher, and the results of individually administered standardized achievement tests. Adaptive behavior-outside school (AB-OS) should be based on information from formal inventories, such as the ABIC (Mercer, 1979a), where appropriate, or informal data collection procedures.

Of particular interest are children who exhibit the pattern of very low intelligence, very low AB-S, and normal AB-OS. Whether these children should be classified and placed in special education programs is a major dilemma currently. Almost by definition such children are "Six Hour Retarded Children"; classifying and placing them in special education programs inevitably leads to the overrepresentation of minority children. In my view, these children *should* be served in special education programs in most instances because, in fact, they have extreme educational needs that typically are beyond the scope of regular classroom resources. The solution of "delabeling" these children does not address their needs. However, the segregated special class for mildly retarded students is an inappropriate solution.

The solution to this dilemma depends on two developments. First, we need a more refined classification system that differentiates between what Mercer (1973) called the "Quasi" and "Comprehensively Retarded." In Mercer's scheme, Comprehensively Retarded persons are those who fail both components of the adaptive behavior dimension and the intelligence dimension. Quasi-Retarded persons exhibit the same pattern except for normal social role performance outside of school. The overrepresentation of minorities in special education classes for the educable mentally retarded in Riverside, California (Mercer, 1973) and other schools largely is attributable to the placement of Quasi-Retarded children with minority backgrounds. Should these children be labeled mentally retarded? Opinions on this issue differ sharply (Goodman, 1979; Mercer, 1979b).

A refinement in the classification system would be beneficial in resolving this dilemma. The terms Comprehensive and Quasi are probably objectionable to many. Use of the term Educational Retardation or some other behaviorally descriptive term of the Quasi-Retarded pattern would be preferable. Greater refinement in the classification system is useful only if there are implications for placement decisions and educational programming. The change suggested may have such implications.

The "Quasi-Retarded" need special services. Special educational services for such children should be oriented toward specific academic need rather than broad social competencies. In most instances the resource program, which includes remedial and compensatory tutorial services, often is a more appropriate option than a self-contained special class. Special class programs for mildly retarded students traditionally place considerable emphasis on broadly defined social competencies and "functional" academic skills (Kolstoe, 1976). This emphasis may be appropriate for comprehensively retarded students but probably is misdirected for most of the "Quasi-Retarded." With few exceptions, Quasi-Retarded students, if placed in special education, should be placed, part-time, in resource programs.

Use of the resource option for Quasi-Retarded students

## Table 4

### Conception of Adaptive Behavior for School-Age Children

### ADAPTIVE BEHAVIOR: SCHOOL BASED

**Rationale:**

1. Mastery of literacy skills is a key developmental task for persons between the ages of 5 and 17.
2. The expectation for and emphasis on educational competencies is common to most if not all major sociocultural groups.

**Assessment:**

1. Collection and consideration of a broad variety of information including teacher interview, review of cumulative records, examination of samples of classroom work, classroom observation, results of group standardized achievement tests, results of individual achievement tests, diagnostic achievement tests, and other informal achievement measures.

### ADAPTIVE BEHAVIOR: OUTSIDE OF SCHOOL

**Rationale:**

1. Mastery of a variety of non-academic competencies also is expected, and a key developmental task between the ages of 5 and 17.
2. The expectation for and opportunities to develop non-academic competencies may vary among sociocultural groups.

**Assessment:**

1. Collection of information on social role performance outside of school in areas such as: peer relations, family relationships, degree of independence, responsibilities assumed, economic/vocational activities, etc.
2. Method of collecting data may include formal measures, interviews with parents, interview with student, etc.

## Table 5

### A Tentative Scheme for Use of Adaptive Behavior Information in Classification and Selection of Program Option

REFERRAL

average

Adaptive Behavior
School Based

Consider other
classifications

significantly
subaverage

averagffe

Intelligence
(Academic Aptitude

significantly
subaverage

Adaptive Behavior
Outside of School

average

significantly
subaverage

"Quasi-retarded"

"Educationally retarded"

"Educationally handicapped"

"Academic aptitude handicap"

Resource Option in
early and middle grades

"Comprehensively retarded"

Mentally Retarded

Mentally Disabled

Special class with
integration
special class

Classification **Classification**

**Selection of Program Option**

**Selection of**
**Program Option**

Selection of
Program Option

alleviates many concerns expressed by Federal District Courts in the placement litigation. The amount of time spent outside the educational mainstream is minimized by the resource option, thus reducing the very proper concern with racial segregation. Placement in the resource room, regardless of the child's classification, also is less stigmatizing.

Refined classification decisions, along with selection of service options (e.g., resource vs. special class), appear to be promising applications of adaptive behavior assessment. Other applications of adaptive behavior data for mildly handicapped students also are promising. General strengths and weaknesses across different domains of behavior may be the initial source of information for developing interventions which are designed to improve social skills, assertiveness, and so on. The information from currently available instruments, such as the ABIC, is not sufficiently precise for direct translation to intervention objectives. Data from the ABIC, AAMD-PS or Vineland can alert us to general needs which, then, can be translated to specific objectives through additional observation and/or interviews.

### Summary

This section, along with the others in this module, discusses the fairness and usefulness of assessment activities, classification decisions, and special education programming. Much remains to be accomplished in the improvement of instruments, expansion of our knowledge base, and application of what we know. Yet, much can be done now to enhance the usefulness and fairness of assessment. We cannot wait for the "system" to change. We can change much of the "system" through what we do. One important question to keep in mind summarizes the intent of nonbiased assessment: Would I be comfortable if my child had been involved in this assessment process?

## SIMULATION 1

Resources needed: Copy of State Special Education Rules and Regulations and/or State Education Code

### Mental Retardation Definition

Analyze the definition of mental retardation used by your state. Is the definition similar to the 1977 Revision of the AAMD *Manual of Terminology and Classification?*

| | Yes | No | Similarities/Differences |
|---|---|---|---|
| 1. Terminology? | | | Comment: |
| mental retardation or? | ____ | ____ | |
| 2. Bi-dimensional? | | | Comment: |
| intelligence | ____ | ____ | |
| adaptive behavior | ____ | ____ | |
| 3. Levels Specified? | | | Comment: |
| mild or educable | ____ | ____ | |
| moderate or trainable | ____ | ____ | |
| severe | ____ | ____ | |
| profound | ____ | ____ | |
| 4. IQ Cut Off Score Specified? | | | Comment: |
| Upper bound— -2 s.d. | ____ | ____ | |
| Mild or educable range | ____ | ____ | |
| 5. Developmental? | | | Comment: |
| age limits | ____ | ____ | |
| 6. Current Status? | ____ | ____ | Comment: |
| 7. Adaptive Behavior Criteria? | | | Comment: |
| Pre-school | ____ | ____ | |
| School age | ____ | ____ | |
| Late adolescence/adult | ____ | ____ | |

## SIMULATION 2

### Clarifying the Meaning of Mild Mental Retardation

Many persons (including students, parents, and teachers) confuse mild mental retardation with the more severe levels of mental retardation. There are three very significant differences between mild and the more severe levels of retardation. Mild mental retardation typically is *not permanent, not comprehensive, and not biologically based.*

Develop a short paragraph that could be used in reports or other written communications to students, parents, and teachers that clarifies the meaning of mild mental retardation.

Compare your statement with statements developed by other participants in the workshop.

Try your statement out with a sample of teachers. Was the statement needed? Did the statement reduce misconceptions about mild mental retardation?

## SIMULATION 3

Resouces needed: Information from a recent case where a child was classified as mildly retarded

### Use of Multifactor Assessment with Mildly Retarded Students

1. Review data from a recent case or a reevaluation of a child placed in a program for the mildly retarded. Is there documentation that each type of data was collected? considered?
2. In your work setting, specify in detail who would collect each type of data, what the data source would be, how the data would be collected, and how each type of data would be considered. (See Table 2.)

   a. **Observational** (in classroom)

   What data are collected?

   Who collects the data?

   How do these data supplement referral information?

   b. **Review of other data** (cumulative records, group test scores, samples of classroom work, etc.)

   What data are collected?

   Who collects these data?

   How are these data used?

   c. **Language Dominance** (determine primary language of child)

   How is primary language determined?

   Who determines primary language?

   How are these data used in subsequent steps in the preplacement evaluation?

   d. **Educational Assessment**

   What *informal* procedures are used to assess educational skills?

   What formal procedures are used to assess educational skills?

   Are both norm-referenced and criterion-referenced measures used?

   Is a comprehensive report developed which describes levels, patterns, and specific strengths and weaknesses in educational skills?

   e. **Sensory-motor** (Correlated Process)

   What measures are used to screen children for visual or auditory processing disabilities?

   If processing difficulties are suggested from the preceding information, who is reponsible for collecting additional information and what additional information is collected?

   How is this information used?

   f. **Adaptive Behavior**

   Who is responsible for collecting adaptive behavior data?

   What measures (formal and informal) are used to assess adaptive behavior data?

   How does the adaptive behavior information influence classification/placement decisions? Selection of program option decisions? Specification of intervention objectives?

   g. **Medical-Developmental History**

   What screening procedures are used to assess medical or developmental history factors?

   Who is responsible for conducting the screening or developmental history?

   If significant problems are suggested in the medical or developmental screening, what further steps are followed to obtain a complete medical evaluation?

h. **Emotional Complication/Behavior Disorders**

What procedures are used to determine if emotional difficulties or behavioral deviations contribute to the problem?

Who is reponsible for assessing these areas? How are these areas assessed?

i. **Intellectual**

How is information from the previous steps used in selecting and interpreting intellectual measures?

Who is responsible for collecting and interpreting data on intellectual performance?

How are data on intellectual performance combined with other information in making classification/placement decisions? in selecting program options?

3. Specify how a mild mental retardation classification/placement decision might be affected depending on the results from each type of data.

4. Describe how the multifactored process is carried out in your setting. Are any steps left out or not carried out adequately? How could *present resources* be used to collect the information required?

## SIMULATION 4

Intelligence test results frequently are misinterpreted. Develop a statement which could be attached to reports to describe the common myths about intelligence. The statement should address the myths that intelligence is unitary, fixed, and predetermined.

After developing the statement, compare your statement with statements developed by other participants in the workshop.

Try your statement out with a sample of teachers. Note their reactions.

## SIMULATION 5

### Selection of an Adaptive Behavior Scale

One of the purposes of adaptive behavior measurement is to improve classification/placement decisions for children and youth with mild mental retardation. Because adaptive behavior measures for this purpose generally are not well developed, careful analysis of available instruments is particularly important. Analysis of what is needed in an ideal instrument can be useful in judiciously applying presently available measures and clarifying the kinds of additional information that should be collected through informal means.

Assume that you anticipate a classification/placement decision. Consider the following questions.

1. What comparison are you making?

2. Do you need a norm-referenced or criterion-referenced instrument to accomplish this purpose? Why?

3. What are the characteristics of an ideal instrument that would provide information on this question? (Hint: You might consider the crucial features of a good norm-referenced test.)

4. What adaptive behavior domains do you want to assess? Why?

5. How well do the instruments in Table 3 meet the criteria you specified in step 2?

## SIMULATION 6

### Collection and Interpretation of Adaptive Behavior Information

1. Collect data on the same child using the ABIC, AAMD-PS, and VSMS. (A child of 7-11 years of age should be selected if possible.)

   Note: The respondent for the ABIC should be the primary caretaker for the child, usually the mother. The respondent for the AAMD-PS is usually the child's teacher. The respondent for the VSMS should be someone who knows the child well, probably the mother in most cases.

2. Compare the data collected with the three instruments. Are the data consistent? How so? Are there inconsistencies? Please specify. Is the information from different respondents consistent or inconsistent?

3. In what ways should the information from each instrument be supplemented? How would you collect the additional information?

4. What effect would the adaptive behavior information for each instrument have on a classification/placement

decision in your setting?

5. To what extent do the different scales provide information on specific educational needs?

## SIMULATION 7

### Analysis of Concepts of Bias

Various methods for identifying possible test bias follow. For each method *predict* the results of research with minority students.

**I. Construct Validity/Content Bias**

**A. Mean Differences Definition**

1. The average scores for black and Hispanic students are lower on the Wechsler scales or the Stanford-Binet than for white children. True False

2. Performance on nonverbal scales are less biased with respect to mean differences. True False

**B. Item Bias or Content Bias Definition**

3. Items on current standardized tests of intelligence are biased against minority students. True False

4. Expert judges representing minority cultures can identify biased items. True False

5. Analysis of performance of minority children on intelligence test items usually supports the judgments of item bias. True False

**C. Psychometric Characteristics Definition**

6. Intelligence test results are less reliable for minority students. True False

7. Relations of items to subtest scores, and subtest scores to scale or over-all IQ scores differ for minority and white children. True False

**D. Factor Analysis Definition**

8. Different factors of intelligence are measured by conventional instruments depending on whether minority or majority children are involved. True False

9. The global or over-all score from conventional tests reflects general intelligence to the same extent for minority and majority students. True False

**II. Atmosphere Bias (Examiner or Situational Effects)**

10. Race of the examiner is a significant influence on the performance of black, Hispanic, and white children on intelligence tests. True False

11. Situational variables, such as incentives, warm vs. cold interaction, etc., have a differential effect on black and Hispanic children. True False

**III. Test Use: Predictive Validity Definition**

12. Conventional IQ tests predict performance on standardized achievement tests equally well for minority and majority students. True False

13. Conventional IQ tests do not predict grades in academic subjects or teacher ratings of academic competence as well for minority students. True False

**IV. Test Use: Social Consequences**

**A. Misuse Definition**

14. Intelligence test results have been misused in justifications of racism and in recommendations for changes in social policy that might restrict rights of minorities. True False

15. Intelligence test results frequently are misinterpreted. True False

**B. Selection Ratios Definition**

16. Use of intelligence test information as part of the assessment process for selection or classification of students frequently results in over- or underrepresentation of minorities True False

17. Elimination of intelligence tests would reduce and, perhaps, eliminate the over- and underrepresentation of minority students in special educational programs. True False

## KEY TO ITEMS AND REFERENCES

| | |
|---|---|
| 1. True | Minorities, particularly economically disadvantaged groups, do, in fact, obtain lower scores on conventional ability tests (Kaufman & Doppelt, 1976). |
| 2. False | Performance or nonverbal tests are not necessarily less biased according to the mean differences criterion. Generally, the performance tests are "less biased" for bilingual groups. |
| 3. ? | See 4 and 5. |
| 4. False | Expert judges can identify items that appear to be biased. However, the agreement among judges typically is quite low (Sandoval, 1979). |
| 5. False | Empirical examinations of item bias rarely yield significant differences among groups (Sandoval, 1979). |
| 6. False | The reliability of intelligence tests is similar for all groups (e.g., Dean, 1977; Sandoval, 1979). |
| 7. False | These relationships remain the same regardless of group membership. |
| 8. False | The factor structure of commonly used instruments remains the same, or highly similar, regardless of group membership (Reschly, 1978). |
| 9. True | The amount of variance attributable to "g" factor is about the same regardless of group membership (Reschly, 1978). |
| 10. False | For both 10 and 11, there is no consistent evidence that atmosphere bias has a significant influence on the performance of any group (Sattler, 1974). |
| 12. True | Again, there is little or no evidence or differential validity from most studies. |
| 13. False | The keyed answer here may be open to some dispute as Goldman and Hartig (1976) reported different correlations depending on group. Most other studies report similar if not identical correlations (Reschly, in press). |
| 14. True | The misuses are well documented by Kamin (1974) and Bersoff (1979). |
| 15. True | See section content on intelligence. |
| 16. True | See placement litigation information in section content or Reschly (1979). |
| 17. Unknown | Many of the plaintiffs in the placement litigation would contend this statement is true. Others, such as Ashurst & Meyers (1973), would argue differently. |

## KEY FOR PRETEST

1. d
2. b
3. False
4. True
5. e
6. False
7. a
8. False
9. d
10. False
11. a
12. False
13. d
14. c
15. False
16 a
17. False
18. c
19. True
20. a
21. c
22. d
23. e
24. False
25. False
26. True
27. b

## REFERENCES

Anastasi, A. *Psychological testing* (4th ed.). New York: MacMillan, 1976.

Ashurst, D., & Meyers, E. Social system and clinical model in school identification of the educable retarded. In R. Eyman, E. Meyers, & G. Tarjan (Eds.), *Sociobehavioral studies in mental retardation.* Washington, D.C.: American Association on Mental Deficiency, 1973.

Baller, W. Charles, D., & Miller, E. Mid-life attainment of the mentally retarded. *Genetic Psychology Monographs,* 1967, *75,* 235-329.

Bersoff, D. Regarding psychologists testily: Legal constraints on psychological assessment in the public schools. *Maryland Law Review,* 1979, *39,* 27-120.

Cegalka, W. Competencies of persons responsible for classification of mentally retarded persons. *Exceptional Children,* 1978, *45,* 26-31.

CORRC (Coordinating Office on Regional Resource Centers) *With bias toward none: Vol. III, A national survey of assessment programs and procedures.* Lexington: University of Kentucky, Undated (distributed in 1979).

Coulter, A., & Morrow, H. *The concept and measurement of adaptive behavior.* New York: Grune & Stratton, 1978.

Cromwell, R., Blashfield, R., & Strauss, J. Criteria for classification systems. In N. Hobbs (Ed.), *Issues in the classification of children.* San Francisco: Jossey-Bass, 1975.

Dean, R. A. Reliability of the WISC-R with Mexican-American children. *Journal of School Psychology,* 1977, *15,* 267-68.

*Diana vs. State Board of Education,* C-70 37 RFP, District Court for Northern California (February, 1970).

Doll, E. *Measurement of social competence.* Circle Pines, MN: American Guidance Service, 1953.

Goldman, R., & Hartig, L. The WISC may not be a valid predictor of school performance for primary-grade minority children. *American Journal of Mental Deficiency,* 1976, *80,* 583-87.

Goodman, J. Is tissue the issue? A critique of SOMPA's models and tests. *School Psychology Digest*, 1979, 8, 47-62.

Grossman, H. (Ed.). *Manual on terminology and classification in mental retardation*. Washington, D.C.: American Association on Mental Deficiency, 1973; 1977.

*Guadalupe vs. Tempe Elementary School District*, 71-435, District Court for Arizona, January, 1972.

Heber, R. A manual on terminology and classification in mental retardation (2nd ed.). *American Journal of Mental Deficiency*, 1961, Monograph Supplement 64.

Jones, R. Protection in evaluation procedures: Criteria and recommendations. In L. Morra (Ed.), *Developing criteria for evaluation of the protection in evaluation procedures provision of Public Law 94-142*. Washington, D.C.: U.S. Office of Education, Bureau of Education for the Handicapped, 1978.

Kamin, L. *The science and politics of IQ*. Potomac, MD: Erlbaum Associates, 1974.

Kaufman, A. A new approach to the interpretation of test scatter on the WISC-R. *Journal of Learning Disabilities*, 1976, 9, 160-68.

Kaufman, A. WISC-R research: Implications for interpretation. *School Psychology Digest*, 1979, 8, 5-27. (a)

Kaufman, A. *Intelligence testing with the WISC-R*. New York: Wiley-Interscience, 1979. (b)

Kaufman, A., & Doppelt, J. Analysis of WISC-R standardization data in terms of stratification variables. *Child Development*, 1976, 47, 165-71.

Kolstoe, O. *Teaching educable mentally retarded children* (2nd ed.). New York: Holt, 1976.

Lambert, N., Windmiller, M., Cole, L., & Figueroa, R. *Manual for the public school version of the AAMD adaptive behavior scale*. Washington, D.C.: American Association on Mental Deficiency, 1975.

*Larry P. et. al. vs. Wilson Riles et. al.* U.S. District Court, Northern District of California, Case No. C-71-2270 RFP, 1972, 1974, 1979.

Leland, H. Theoretical considerations of adaptive behavior. In A. Coulter & H. Morrow (Eds.), *The concept and measurement of adaptive behavior*. New York: Grune & Stratton, 1978.

MacMillan, D. *Mental retardation in school and society*. Boston: Little, Brown, 1977.

MacMillan, D., Jones, R., & Aloia, G. The mentally retarded label: A theoretical analysis and review of research. *American Journal of Mental Deficiency*, 1974, 79, 241-61.

Matarazzo, D. *Wechsler's measurement and appraisal of adult intelligence (5th and enlarged edition)*. Baltimore: Williams & Wilkins, 1972.

*Mattie, T., et. al. vs. Charles E. Holladay, et. al.* C.A. No. DC-75-31-S. U.S. District Court for the Northern District of Mississippi, 1979.

Mercer, J. *Labeling the mentally retarded*. Berkeley, CA: University of California Press, 1973.

Mercer, J. Protection in evaluation procedures. In L. Morra (Ed.), *Developing criteria for evaluation of the protection in evaluation procedures provision of Public Law 94-142*. Washington, D.C.: U.S. Office of Education, Bureau of Education for the Handicapped, 1978.

Mercer, J. *Technical Manual: SOMPA: System of Multicultural Pluralistic Assessment*. New York: Psychological Corporation, 1979. (a)

Mercer, J. In defense of racially and culturally nondiscriminatory assessment. *School Psychology Digest*, 1979, 8, 89-115. (b)

Morra, L. (Ed.). *Developing criteria for evaluation of the protection in evaluation procedures provision of Public Law 94-142*. Washington, D.C.: U.S. Office of Education, Bureau of Education for the Handicapped, 1978.

Newland, T. E. Psychological assessment of exceptional children and youth. In W. Cruickshank (Ed.), *Psychology of exceptional children and youth (4th ed.)*. Englewood Cliffs, N.J.: Prentice Hall, 1980.

Oakland, T., & Goldwater, D. Assessment and interventions for mildly retarded and learning disabled children. In G. Phye & D. Reschly (Eds.), *School psychology: Perspectives and issues*. New York: Academic Press, 1979.

Page, E. Tests and decisions for the handicapped. In L. Morra (Ed.), *Developing criteria for evaluation of the protection in evaluation procedures provisions of Public Law 94-142*. Washington, D.C.: U.S. Office of Education, Bureau of Education for the Handicapped, 1978.

Reschly, D. WISC-R factor structures among Anglos, Blacks, Chicanos, and Native American Papagos. *Journal of Consulting and Clinical Psychology*, 1978, 46, 417-22.

Reschly, D. Nonbiased assessment. In G. Phye & D. Reschly (Eds.), *School psychology: Perspectives and issues*. New York: Academic Press, 1979.

Reschly, D. Concepts of bias in assessment and WISC-R research with minorities. In A. Engin (Ed.), *WISC-R Handbook*. Washington, D.C.: National Association of School Psychologists, in press.

Reschly, D., & Davis, R. Comparability of WISC and WISC-R scores among borderline and mildly retarded children. *Journal of Clinical Psychology*, 1977, 33, 1045-48.

Reschly, D., & Reschly, J. Predictive utility of WISC-R factor scores for four groups. *Journal of School Psychology*, 1979, 17, 355-61.

Richmond, B., & Kicklighter, R. *Children's Adaptive Behavior Scale*. Atlanta: Humanics Limited, 1980.

Robinson, N., & Robinson, H. *The mentally retarded child* (2nd ed.). New York: McGraw-Hill, 1976.

Sandoval, J. The WISC-R and internal evidence of test bias with minority groups. *Journal of Consulting and Clinical Psychology*, 1979, 47, 919-27.

Salvia, J. & Ysseldyke, J. *Assessment in special and remedial education*. Boston: Houghton-Mifflin, 1978.

Sattler, J. *Assessment of children's intelligence*. Philadelphia: Saunders, 1974.

Silverstein, A. WISC subtest patterns of retardates. *Psychological Reports*, 1968, 23, 1061-62.

Tucker, J. Operationalizing the diagnostic-intervention process. In T. Oakland (Ed.), *Psychological and educational assessment of minority children*. New York: Brunner/Mazel, 1977.

Turnbull, H. The past and future impact of court decisions in special education. *Kappan*, 1978, 60, 523-27.

Ysseldyke, J. Implementing the "Protection in Evaluation Procedures" provisions of P.L. 94-142. In L. Morra (Ed.), *Developing criteria for evaluation of the protection in evaluation procedures provisions of Public Law 94-142*. Washington, D.C.: U.S. Office of Education, Bureau of Education for the Handicapped, 1978.

Zigler, E. Familial mental retardation: A continuing dilemma. *Science*, 1967, 155, 292-98.

## RECOMMENDED READING

The following sources are highly recommended. They provide basic information on mental retardation, intellectual assessment, adaptive behavior, and nonbiased assessment of minorities.

1. **Basic Sources in Mental Retardation**

MacMillian, D. *Mental retardation in school and society* (2nd ed.). Boston: Little, Brown, 1977.

Robinson, N., & Robinson, H. *The mentally retarded child* (2nd ed.) New York: McGraw-Hill, 1976.

**Comment:** Both books are excellent sources of information on mental retardation. The Robinson & Robinson book is particularly good in reviewing basic areas of psychology and then applying the information to mental retardation. The MacMillan book provides more information on mildly retarded persons and on educational programming.

2. **Adaptive Behavior**

Coulter, A., & Morrow, H. (Eds.). *The concept and measurement of adaptive behavior*. New York: Grune & Stratton, 1978.

**Comment:** Leading scholars in the field contributed chapters to the Coulter & Morrow book; it is, perhaps, the best current source on adaptive behavior. The book is an accurate reflection of the current literature in that diverse points of view are expressed and more questions are raised than answers given.

3. **Intelligence**

Sattler, J. *Assessment of children's intelligence*. Philadelphia: Saunders, 1974.
Note: A new edition of Sattler's book is scheduled for publication in 1981.

**Comment:** The Sattler text provides a comprehensive treatment of concepts of, research on, and measures of intelligence.

4. **Nonbiased Assessment**

Oakland, T. (Ed.). *Psychological and educational assessment of minority children*. New York: Brunner/Mazel, 1977.

Reschly, D. Nonbiased assessment. In G. Phye & D. Reschly (Eds.), *School psychology: Perspectives and issues*. New York: Academic Press, 1979.

Mercer, J. *Technical Manual: SOMPA*. New York: Psychological Corporation, 1979.

**Comment:** The preceding sources present different perspectives on the elimination of bias in assessment. Oakland provides chapters on historical antecedents, current legislation, use of tests, diagnostic-intervention programs, and multifactor assessment, and extensive appendices with information on ethical standards, language dominance measures, etc. Reschly discusses bias in assessment from the perspective of different definitions of bias; underlying assumptions of the courts in placement bias cases; research on bias in tests; and procedures for enhancing fairness in assessment. Mercer's approach to eliminating bias stresses the collection of a broad variety of data; different models of assessment; and use of pluralistic norms.

## Appendix 1

### Protection in Evaluation Procedures Provisions of PL 94-142 Rules and Regulations

Protection in Evaluation Procedures

§ 121a.530 General.

(a) Each State educational agency shall insure that each public agency establishes and implements procedures which meet the requirements of §§ 121.a-530-121a.534.

(b) Testing and evaluation materials and procedures used for the purposes of evaluation and placement of handicapped children must be selected and administered so as not to be racially or culturally discriminatory.

(20 U.S.C. 1412(5)(C).)

§ 121a.531 Preplacement evaluation.

Before any action is taken with respect to the initial placement of a handicapped child in a special education program, a full and individual evaluation of the child's educational needs must be conducted in accordance with the requirements of § 121a.532.

(20 U.S.C. 1412(5)(C).)

§ 121a.532 Evaluation procedures.

State and local educational agencies shall insure, at a minimum, that:

(a) Tests and other evaluation materials:

(1) Are provided and administered in the child's native language or other mode of communication, unless it is clearly not feasible to do so;

(2) Have been validated for the specific purpose for which they are used; and

(3) Are administered by trained personnel in conformance with the instructions provided by their producer;

(b) Tests and other evaluation materials include those tailored to assess specific areas of educational need and not merely those which are designed to provide a single general intelligence quotient;

(c) Tests are selected and administered so as best to ensure that when a test is administered to a child with impaired . . . sensory, manual, or speaking skills, the test results accurately reflect the child's aptitude or achievement level or whatever other factors the test purports to measure, rather than reflecting the child's impaired sensory, manual, or speaking skills (except where those skills are the factors which the test purports to measure);

(d) No single procedure is used as the sole criterion for determining an appropriate educational program for a child; and

(e) The evaluation is made by a multidisciplinary team or group of persons, including at least one teacher or other specialist with knowledge in the area of suspected disability.

(f) The child is assessed in all areas related to the suspected disability, including, where appropriate, health, vision, hearing, social and emotional status, general intelligence, academic performance, communicative status, and motor abilities.

(20 U.S.C. 1412(5)(C).)

Comment. Children who have a speech impairment as their primary handicap may not need a complete battery of assessments (e.g., psychological, physical, or adaptive behavior). However, a qualified speech-language pathologist would (1) evaluate each speech impaired child using procedures that are appropriate for the diagnosis and appraisal of speech and language disorders, and (2) where necessary, make referrals for additional assessments needed to make an appropriate placement decision.

### § 121a.533 Placement procedures

(a) In interpreting evaluation data and in making placement decisions, each agency shall:

(1) Draw upon information from a variety of sources, including aptitude and achievement tests, teacher recommendation, physical condition, social or cultural background, and adaptive behavior;

(2) Insure that information obtained from all of these sources is documented and carefully considered;

(3) Insure that the placement decision is made by a group of persons, including persons knowledgeable about the child, the meaning of the evaluation data, and the placement options; and

(4) Insure that the placement decision is made in conformity with the least restrictive environment rules in §§ 121a.550-121a.554.

(b) If a determination is made that a child is handicapped and needs special education and related services, an individualized education program must be developed for the child in accordance with §§ 121a.340-121a.349 of Subpart C.

(20 U.S.C. 1412(5)(C); 1414(a)(5).)

Comment: Paragraph (a)(1) includes a list of examples of sources that may be used by a public agency in making placement decisions. The agency would not have to use all the sources in every instance. The point of the requirement is to insure that more than one source is used in interpreting evaluation data and in making placement decisions. For example, while all of the named sources would have to be used for a child whose suspected disability is mental retardation, they would not be necessary for certain other handicapped children, such as a child who has a severe articulation disorder as his primary handicap. For such a child, the speech-language pathologist, in complying with the multisource requirement, might use (1) a standardized test of articulation, and (2) observation of the child's articulation behavior in conventional speech.

### § 121a.534 Reevaluation.

Each State and local educational agency shall insure:

(a) That each handicapped child's individualized education program is reviewed in accordance with §§ 121a.340-121a.349 of Subpart C, and

(b) That an evaluation of the child, based on procedures which meet the requirements under § 121a.532, is conducted every three years or more frequently if conditions warrant or if the child's parents or teacher requests an evaluation.

(20 U.S.C. 1412(5)(C).)

# NONBIASED ASSESSMENT OF EMOTIONALLY DISTURBED STUDENTS

Martin L. Tombari
Austin, Texas

## Table of Contents

Introduction .......... 123

Objectives .......... 123

Pretest .......... 124

Self-Assessment of Current Practices .......... 124

Assessment of Emotional Disturbance: General Considerations .......... 125
  Defining Emotional Disturbance .......... 125
  Testing vs. Assessment .......... 126
  Factors Affecting Judgments About Emotional Disturbance .......... 126
  Traditional Data Bases for Assessment .......... 126
  Behavioral Assessment as an Alternative .......... 128
  Behavioral Assessment of Minority Students .......... 128

Assessment of Emotional Disturbance: Specific Guidelines .......... 130
  Data for Decision Making .......... 130
  Classification Decisions .......... 130
  Programming Decisions .......... 133
  Collecting Evaluation Data .......... 136
  Remaining Questions About Behavioral Assessment .......... 136

Conclusion .......... 137

Posttest .......... 138

Simulations .......... 139

Key to Pretest .......... 139

Key to Posttest .......... 139

References .......... 139

Recommended Readings .......... 140

## Assessing Emotionally Disturbed Students: An Introduction

The underlying assumption of this section is that assessment of emotionally disturbed students by school psychologists must be related directly to educational decision making. Because the statement is deceptively simple, one may be tempted to respond, "All my assessments are relevant to educational decision making. As a *school* psychologist, I evaluate children to help them in schools." In practice, many school psychologists assess emotionally disturbed children and youth and hope that educators will use the report. Others correctly prefer to use the decisions that educators must make about emotionally disturbed students to guide their assessments.

Education personnel often find the assessments of school psychologists to be irrelevant in daily classroom decisions (Bardon, 1971). We can understand this complaint better if we inspect the decisions educators make when they plan and carry out programs for emotionally disturbed (E.D.) students. Educators depend on school psychologists to help them decide the specific objectives of a program, the techniques for teaching these objectives, and the most appropriate methods for evaluating their attainment. Unfortunately, the standard assessments of E.D. students by school psychologists often are not aimed at providing the data to make these decisions. Rather, their primary function is (and has been) to provide clinically rich descriptive data from which hypotheses can be generated about the nature and cause of emotional disturbance. This information may help to label a student but rarely meets the needs of educators who must plan programs which help that student.

The objectives of school psychologists who assess students for emotional disturbance differ from those of a clinical psychologist or psychiatrist. The latter mental health professionals usually treat clients in private offices, clinics, or hospitals. Their assessments are designed to help them to make decisions about treatments which they will carry out in these settings. School psychologists, on the other hand, rarely work directly with emotionally disturbed students, either alone or in non-school settings. Rather, in school settings, it is the educators who work directly with students.

School psychologists, then, must recognize that their assessments are relevant to the following realities: (a) Emotionally disturbed students will be helped primarily in a school setting; (b) programs for them will be developed and carried out by educators; (c) the objectives of such programs must be stated in terms of measurable changes in attitudes, beliefs, and behaviors; and (d) educators look to school psychologists for guidance in deciding which objectives to teach, how to teach them, and how to evaluate progress.

This section of the module is concerned with the nonbiased assessment of emotionally disturbed students. The position taken is that biased assessments are those that place students in ineffective programs that do not provide appropriate opportunities for social and educational development. Nonbiased assessments, on the other hand, lead to effective programs that help emotionally disturbed students to achieve socially and educationally important goals and objectives. When assessments are related directly to educational decision making, they are potentially less biased than those that are not so related.

Assessments of emotional disturbance are enhanced and more information is acquired through direct observation, systematic interviewing, and self-reports. The use of projective devices is not ruled out but their relevance to educational decision making is questionable. Consequently, the assessment of emotional disturbance through figure drawings, TAT, CAT, Rorschach, or sentence completion techniques is not discussed in this paper.

### Objectives of Section

1. To recognize or formulate your own beliefs of and assumptions on the nature of emotional disturbance.
2. To distinguish between psychometric testing and psychological assessment.
3. To list factors that affect one's judgment to label a student emotionally disturbed.
4. To relate assessment needs to PL 94-142 requirements for the development and implementation of IEPs for exceptional children.
5. To describe the essential differences between behavioral assessments and more traditional assessments of emotional disturbance.
6. To discuss the importance of relationships between assessment and educational programming for emotionally disturbed children.
7. To relate the preceding objective to issues of nonbiased assessment.
8. To evaluate opinions on the bias of behavioral assessment against minority groups.
9. To state the three basic decisions that assessments of emotional disturbance must help to make.
10. To list the sources of data for each of the three decisions.
11. To describe assessment techniques that help to gather the information for the preceding decisions.
12. To describe three sources of data that help to decide if a student qualifies for the emotionally disturbed label.
13. To describe how an interview with a teacher can help to gather data to facilitate decisions on which educational strategies best can help an emotionally disturbed child.
14. To summarize the evidence for the validity and reliability of classroom observations, behavioral interviews, and self-reports for assessing emotional disturbance.

## PRETEST

1. Educators and mental health professionals alike are concerned about the critical indicators of emotional disturbance. Which statement below best indicates the correct state of affairs concerning indicators of ED?
   a. The best indicators of emotional disturbance are conduct disorders seen in elementary school.
   b. There is basic agreement among mental health professionals about the critical indicators of ED.
   c. There is much disagreement among mental health professionals about the critical indicators of ED.
   d. Teachers and school psychologists view the indicators of ED similarly.

2. Most experts agree that the causes of ED lie in early childhood rearing patterns.

   True
   False

3. Which of the following statements represents the best comparison between psychometric testing and psychological assessment of ED?
   a. Psychological assessment is a more objective, data-based process.
   b. Both are subjective processes using essentially subjective data.
   c. Psychometric testing provides the data for the subjective judgments made in psychological assessment.
   d. They are essentially the same.

4. Assessment of emotional disturbance is influenced by the values of particular communities and individuals in the community.

   True
   False

5. Nonbehavioral assessments of emotional disturbance tend to
   a. be concerned with variables unrelated to labeling decisions
   b. be unrelated to educational decisions
   c. provide more valid data than behavioral assessments on which to base labeling decisions
   d. depend heavily on interviews and self-reports

6. Self-reports of thoughts and feelings are minimized in behavioral assessment.

   True
   False

7. Closer relationships between assessments of emotional disturbance and educational interventions to help the emotionally disturbed will lead to
   a. more effective programs
   b. more nonbiased assessment
   c. better meeting the mandates of PL 94-142
   d. a & c only
   e. a, b, & c

8. The strength of behavioral assessment lies in its close relationship with educational programming.

   True
   False

9. The decision to label a child emotionally disturbed should be based on data and affirmative answers to which of the following question(s):
   a. Do the child's problems make him appear significantly different from his peers or classmates?
   b. Do the child's problems interfere with his learning and development?
   c. Are the problems such that the child can't be helped in the regular classroom?
   d. all of the above
   e. a & c only

10. In behavioral assessment, normative comparisons are not drawn between students.

   True
   False

### Self-Assessment of Current Practices

Place a check next to those statements that describe your techniques for assessing emotional disturbance.

______ 1. Identify specific educational decisions that need to be made about a referred student prior to beginning the assessment process.

______ 2. Select techniques that will provide data to help make the preceding decisions.

______ 3. Gather observable descriptions of the student's problems from relevant school personnel.

______ 4. Identify specific educational goals for the student.

______ 5. Focus assessment on the specific problems and goals identified previously which will form the objectives of the student's IEP.

______ 6. Gather observational data on the student in various school settings where different tasks are required of the student and different structures are given to her/him.

______ 7. Gather data on the extent of similar problems among classmates.

______ 8. Gather data on the details and success of efforts to help the student in the regular classroom.

______ 9. Use a variety of assessment techniques, including behavioral observations, systematic interviews, and self-reports.

______ 10. Gather data on how the tasks, structures, and motivations of different environments affect the student's behavior.

______ 11. Gather criterion-referenced data on academic skills deficits.

______ 12. Provide recommendations to increase the student's appropriate behaviors and decrease the inappropriate ones.

______ 13. Provide recommendations to help the student to achieve IEP objectives. Included should be tasks that provide success, the structure that allows the student to work best, and the motivation techniques likely to help the student to initiate, persist in, and complete program tasks.

______ 14. Provide suggestions on how to evaluate the student's achievement of IEP objectives.

## Assessment of Emotional Disturbance: General Considerations

### Defining Emotional Disturbance

The label "emotionally disturbed" perhaps is the most problematic of all special education labels. Psychologists, educators, and parents shrink at the prospect of applying this label to students because of the surplus of meanings attached to it. Moreover, the problem is compounded because there is little agreement about what the label denotes. Psychiatrists, school psychologists, classroom teachers, psychiatric social workers, and other mental health professionals differ on what the term means, what are the salient symptoms of emotional disturbance, and what is its etiology.

---

*Task 1*

*Examine your beliefs and assumptions about the nature of emotional disturbance by answering the following questions:*

*A. In your opinion, what are the signs of emotional disturbance? List at least seven of them.*

*B. What do you consider to be the major causes of emotional disturbance?*

---

Fremont and his associates asked questions like these of teachers, school psychologists, and other mental health workers (Fremont, Kingsporn, & Wilson, 1976). The responses they received were classified into two groups: conduct disorders and personality disorders. Conduct disorders are behavioral problems (e.g., attention seeking, general classroom disruptiveness, fighting, hyperactivity, profanity, uncooperativeness, and temper tantrums). Personality disorders subsume characteristics such as anxiety, depression, hypersensitivity, lack of self-confidence, disinterestedness, shyness, withdrawal, and phobias. Conduct disorders encompass more observable, measurable, situation-specific problems whereas personality disorders are more abstract and inferential.

School personnel, such as teachers and counselors, are more likely to describe emotional disturbance in terms of conduct disorders. Psychiatrists, clinical psychologists, and psychiatric social workers typically describe emotional disturbance as a personality disorder and school psychologists tend to share their views.

The issue is not which group is right or wrong but, rather, which position leads to educational efforts that are likely to benefit emotionally disturbed students in the least restrictive school setting. The manner in which educators perceive and conceptualize student problems affects their expectations for and actual teaching behavior toward students. Teachers who conceptualize students' problems in more behavioral terms (as opposed to medical model concepts) hold high expectations for the usefulness of classroom efforts (Tombari & Bergan, 1978). Also, teachers are likely to provide more extensive remedial efforts to help students when problems are defined in behavioral rather than personality or medical model terminology (Bergan, Byrness, & Kratochwill, 1979).

The second question of Task 1 asked your conception of the major causes of emotional disturbance. Psychiatrists, school psychologists, and other mental health professionals typically view the causes in terms of an underlying mental illness. This illness often is believed to be an outgrowth of early childhood rearing patterns. For years, many special educators viewed certain emotionally disturbed students as minimally brain injured (Hewett, 1968). There is little research to support either viewpoint. In fact, little is known about the causes of emotional disturbance.

The issue centers on which conception of emotional disturbance is more likely to bring about effective school programs for needy students. The opinion of this author is that a behavioral rather than a medical model is of greater value to school programs for emotionally disturbed students.

Imagine yourself a teacher of emotionally disturbed students, or a regular education teacher with an E.D. student mainstreamed in your classroom. You are told that the cause of the student's problems is some mental illness or minimal brain injury. Your view of the relationship between remedial classroom efforts and the "real cause" of the emotional disturbance very likely could be expressed in the often heard lament, "What can we do at school to really help this student? The damage already has been done!"

Now consider an alternative view of emotional disturbance. This view states that students' behaviors are learned, that they reflect how people behave toward them, particularly teachers and parents. If adults and children learn improved ways of dealing with these behaviors, the problems are more likely to be resolved. This perspective, in the opinion of this author, more likely leads to patient, consistent, and effective interventions at home and school. Thus, the viewpoint in this paper is that emotional disturbance is defined best as inappropriate overt or covert behavior (e.g., thoughts, fears, attitudes). The immediate environment of the student significantly contributes to the development and maintenance of this behavior. Likewise, changes in the environment can help the student to learn appropriate ways of behaving. This view facilitates communication between parents, teachers, and school psychologists. The behavioral perspective on the nature of emotional disturbance and the factors influencing its development provides the rationale for the assessment procedures.

### Testing vs. Assessment

---

*Task 2*

*Before beginning this section on the assessment of emotional disturbance, describe in your own words the important differences between psychometric testing and psychological assessment.*

---

Psychometric testing is the administration and scoring of tests (e.g., intelligence, personality, and achievement tests). It is largely an objective procedure. Computing an IQ, figuring a percentile, or drawing a profile on a personality inventory are straightforward tasks requiring little subjective judgment. However, when we make interpretations or draw conclusions from the data (e.g., that a student is an overachiever, unintelligent, or obsessive-compulsive), we cross the boundary of testing into the realm of assessment.

Assessment is the interpretation of data derived from psychometric testing and other information-gathering techniques. Testing techniques provide the raw data from which assessments are formed. Assessment is a judgmental process based on data. When one assesses intelligence, one makes one's best judgment of a student's ability to adapt to the environment; the IQ score assists in making that judgment. Similarly, in assessing emotional disturbance, school psychologists make a judgment that a student cannot adjust to normal classroom rules and routines and requires a special education program to do so. Observation, interview, and self-report data help one to make this judgment.

School psychologists often lose sight of this distinction between testing and assessment. They would like to believe that their judgments are objective because they use objective data-gathering techniques. However, assessments are subjective and prone to all the biases inherent in any judgmental activity.

Thus, school psychologists must realize that when they call someone "emotionally disturbed" they are making a subjective assessment that represents their best judgment. This is not to imply that the data base for such judgments is subjective. Quite the contrary. Some data collection strategies are objective, reliable, and valid. However, after all the data are collected, there is no magic formula for analysis that insures correct labeling, classification, and intervention.

### Factors Affecting Judgments About Emotional Disturbance

---

*Task 3*

*Given that the decision to label a student emotionally disturbed is subjective, list those factors that may affect your judgment in so labeling a student.*

---

In this section we examine the factors that affect our assessments of emotional disturbance, the traditional data base for such judgments, and a data base that is consistent with the requirements of PL 94-142 and the thesis of this paper.

Many factors influence our judgments of whether a student is emotionally disturbed. One factor is our choice of aspects of a student's behavior to observe. We know that we can increase the reliability and validity of our judgments by increasing the number of behaviors we observe and the length of time over which our observations are conducted; however, time and money require that we observe selectively. We can observe samples of actual behavior; obtain self-reports of behavior, beliefs, attitudes, or fears; or examine the reports of others. The techniques we select inevitably affect the pictures we obtain of students and the judgments we make of them.

The setting in which we observe a student influences our assessments. Many school psychologists see children only in a one-to-one testing situation. Observing them on playgrounds, in reading or physical education classes, cafeterias, or at home certainly provides a different perspective.

The occasions on which we observe also influence our picture of a student. The data obtained by observing, interviewing, or testing a child on Monday morning are different from those obtained Friday afternoon. Likewise, a student's behavior in class, on tests, or in interviews will vary from early morning, to just before lunch, or after physical education class.

The social desirability of a behavior pattern also can affect judgments about a child. The behaviors that are judged acceptable vary from decade to decade, by geographical region, and from one individual to another. Sex role expectations and cultural or racial stereotyping may bias our conclusions about the appropriateness of behaviors observed. (The section of this module by Henderson, "Nonbiased Assessment: Sociocultural Considerations, further discusses this issue).

Begelman (1975), in a thoughtful analysis of classification and labeling issues, cautioned that identifying problem behaviors is not a value-neutral task:

> The very notion of a problem behavior is conceptually linked to issues in the realm of law, ethics, and value. . . . To call someone a behavior problem is not to depict it in the abstract, but to judge it concretely, typically in terms of a tacit system of values. (Begelman, 1975, p. 163)

Thus, the what, when, where, and how of data collection affect our judgments. We have no completely objective criteria for defining emotional disturbance. This problem may lead some school psychologists to avoid participation in E.D. assessments or labeling a child as emotionally disturbed. However, one's appreciation of the difficulties of applying a label should not be a deterrent; rather, it should be an impetus for making the assessment as comprehensive and valid as possible.

### Traditional Data Bases for Assessment

Specific guidelines and criteria exist for assessing mental retardation or even learning disabilities, but comparable guidelines do not exist for assessing emotional disturbance. However, examining the psychological reports of such assessments allows us to determine the data base that is traditionally used by school psychologists to identify emotional disturbance. These reports usually include psychologists' observations of the students at the time of testing, IQ test, norm-referenced achievement test, human figure drawing, and other projective techniques, such as the TAT, CAT, or Rorschach. Also included are the anecdotal reports by teachers and parents. An example of an E.D. evaluation by a school psychologist using these data sources follows.

## Psychological Evaluation

Name: Bill Daniels
D.O.B.: 2/1/68
Dates of Testing: 10/5, 10/7, 10/13, 10/19
Examiner: Robert Thomas

**Referral:**

Bill, a 12-year-old male referred for evaluation by his J.H.S. principal, is on school suspension for bringing a knife to school. His teachers report lack of interest in all school subjects, difficulty making friends, and immature behavior. His parents report that Bill lately has been spending a lot of time in his room by himself. They also suspect him of drinking alcohol from the liquor closet.

**Bases for Evaluation:**

Clinical
WISC-R
MMPI
H-T-P Technique
Rorschach
TAT
WRAT
Teacher observations

**Behavioral Observations:**

Bill is an attractive, well-built, 12-year-old who was dressed neatly and appropriately for the assessment sessions. He remained cooperative and motivated throughout. He appeared anxious; nevertheless, he persisted at tasks, and spontaneously volunteered information about himself and his family. Bill expressed concern with his inability to control events—almost as if fate were making him bad. He finds school boring and friends difficult to make. He said that he brought the knife to school because another boy threatened to get him. Bill said that you should believe someone who says that he is going to get you in this area of the city. He stated that he felt depressed, tense, and unable to determine what was happening to his life.

**WISC-R:**

VIQ = 90
PIQ = 95
FSIQ = 94

Bill's FSIQ on the WISC-R placed him in the average range of intellectual functioning. Bill's scores do not correspond with his school performance, with grades ranging from C to F. The test results suggest that emotional factors and/or perceptual motor difficulties may be interfering in Bill's school performance. The subtest scores also suggest the presence of anxiety and depressive tendencies.

**MMPI**

The validity scales indicate this profile to be valid. The clinical profile indicates the presence of nonconformist and deviant attitudes and values, as well as aggressive and antisocial tendencies. The tendency to act out impulsively in response to anxiety and depression also is indicated. There appears to be some concern with sexuality which is common for a boy this age. Tendencies to use repression and denial to defend against psychological conflicts are also suggested. The profile, in addition, indicates current moderate depression, pessimism, and worry.

**Projective Results:**

Bill's responses to the HTP, Rorschach, and TAT suggest pervasive feelings of tension and anxiety which accompany feelings of uncertainty, insecurity, and lack of control over events. Emotional ability, impulsivity, and the tendency to act out emotional responses also are suggested. There appears to be much conflict surrounding the expression of aggressive or hostile impulses. Concerns with sexuality also are suggested. Bill copes with these conflicts through withdrawal, escape into fantasy, denial, rationalization, and acting out.

**Achievement Tests:**

Bill has a grade equivalent score of 4.2 on the WRAT reading, 4.4 in math; all scores almost two years below grade level. Given Bill's IQ, these scores suggest that emotional problems are interfering with school learning.

**Teachers Observations:**

Teachers report that Bill is uncooperative, rarely turns in assignments, often is late to class, and does not relate to adults. He has few friends. The friends he has are "misfits" like Bill.

**Summary:**

Bill obtained a FSIQ of 94, placing him in the average range of intellectual functioning. His school achievement is below what would be expected given his IQ. This suggests the emotional difficulties may be affecting his performance at school. Bill appears to be experiencing general feelings of anxiety, depression, and loss of control. The testing indicates the presence of pervasive feelings of uncertainty, inadequacy, and helplessness. Bill appears to be emotionally labile with tendencies to act impulsively in response to feelings of anxiety and depression. The testing also indicates aggressive, oppositional, and antisocial tendencies, along with conflicts surrounding the expression of these impulses. Much of Bill's current distress appears to be related to conflicts surrounding feelings involving himself and relationships with his parents. It appears that the knife incident was an impulsive act in response to a threat and does not necessarily indicate a pattern of delinquent behavior.

**Recommendations:**

1. Placement in a special education program classified as emotionally disturbed.
2. Individual psychotherapy to enable Bill to explore current discomfort and to aid in the alleviation of his current anxiety and depression.
3. Family therapy, at a later date, if assessed to be appropriate, to improve family relationships.

In light of the preceding discussion on the need for clear relationships between assessment and school intervention, and for an objective data base for our subjective judgments, numerous questions are posed by this assessment.

**Specifying the problem.** What is the school problem for which Bill is being judged emotionally disturbed? The report itself rules out the knife incident as significant. The only other specific behaviors mentioned are lateness and missed assignments, but no reliable measures are presented on the extent of either problem. Thus, with respect to behavior problems alone, there appears little basis for labeling Bill E.D.

The assessment turned up numerous, hypothesized, affective problems that include anxiety, depression, sense of loss of control, feelings of inadequacy, and helplessness. Outside of the testing situation, how is each problem manifested? Nothing is said about the specific nature of the anxiety or depression. The projective techniques are a rich source of hypotheses, none of which is validated in the assessment. This report presents a weak case for a diagnosis of emotional disturbance. Although the projective techniques suggest tendencies to act out hostile impulses and antisocial, oppositional, and aggressive feelings, no incidents of fighting, classroom disruption, or widespread noncompliance with school rules are reported. It appears that Bill has learned to cope with these hypothesized tendencies, although the report chooses to ignore the possibility of such an asset.

**Setting goals and objectives for the E.D. program.** The assessment is rich in clinical descriptions of hypothesized problems but it provides little help to educators who must decide on goals and objectives for Bill's E.D. program. Which behavioral change should the program focus on: social skills? handing in assignments? conversing with adults? getting to school on time? verbal acting out of hostile impulses? What affective change should the teacher try to bring about: positive self-talk? skills in coping with anxiety? expressions both overt and covert, of confidence in oneself? What healthy beliefs should the program try to foster: the value of school learning? importance of friends? What academic skills should be taught? At the very least, an assessment should assist educators in specifying the foci of an E.D. program.

**Suggestions for achieving goals and objectives.** Educators look to psychologists for clues on how to help children. The report on Bill offers none. Moreover, the report implies that the cause of the problem lies in disturbed family relationshps. The obvious conclusion is, "Until the family problems are resolved, what success can we have at school?"

**Suggestions for evaluating progress in the E.D. program.** How will educators know if efforts to help Bill are successful? The report ignores the issue of evaluating Bill's progress.

This report is by no means a caricature. It is typical of those read in a survey of psychological reports (Tombari & Brein, 1978). The purpose of such reports is not to meet the needs of educators for program decisions but to meet the need of psychologists for classification and labeling decisions—a legitimate need that grows out of legal mandates. However, the strategies suggested in this paper allow psychologists to conduct assessments for emotional disturbance that meet both legal and educational needs.

**What type of data should an E.D. assessment include?** Answers to this question come, in part, from PL 94-142's requirements for individualized educational programs (IEPs). This law requires schools to develop, carry out, and evaluate an IEP for every child receiving special service. Such a plan must include information on the following:

- Current levels of performance in academic and social-emotional areas.
- Previous educational efforts to help the student.
- General areas of need.
- Instructional objectives.
- The nature of services to help the student achieve the objectives.
- The expected duration of services.
- Procedures for measuring student progress.

To the greatest extent possible, an assessment of emotional disturbance should provide data that facilitate making decisions about each IEP element. The following discussion of behavioral assessment outlines ways that are likely to facilitate decisions about labeling, IEP development, and evaluation.

### Behavioral Assessment As an Alternative to Traditional Assessments of Emotional Disturbance

**Defining Behavioral Assessment.** The primary functions of behavioral assessments are to identify problems in measurable terms, isolate the environmental variables that control the problems, select a treatment technique to resolve the problem, and evaluate the effects of the treatment (Ciminero, Calhoun, & Adams, 1977).

In behavioral assessment the focus of measurement is on the problem behaviors which are reported by teachers, parents, or the students themselves. The behaviors that are measured may be overt (e.g., fighting, swearing, out-of-seat, and general disruptiveness), covert and cognitive (e.g., beliefs, images, and thoughts), or covert and affective (e.g., attitudes, fears, feelings). Many naive behaviorists and mental health professionals incorrectly assume that behavioral assessment is confined to studying operant behavior in natural settings. Behavioral assessment, however, is concerned with what students *say* they do, think, and feel as well as with what they actually do, think, and feel.

Behavioral assessment makes no assumptions about the underlying meaning of observed behaviors and reported thoughts or feelings. These variables become the very objects of assessment and intervention. Consequently, they are studied directly in the environment in which they occur. If a student is referred for the evaluation of emotional disturbance because he fights, has no friends, thinks adults are all out to get him, and reports frequent headaches and vomiting, then the objects of assessment and program planning become the same fighting, social skills, irrational thoughts, and physical states which are examined, as much as possible, in the environment in which they occur.

The contrast between traditional assessment and behavioral assessment is evident. In the first, the importance of behavioral topography and setting is down played. Reported behaviors, thoughts, feelings, or fears are regarded as symptoms of some underlying disturbance. Inasmuch as such assessment assumes that stable traits within the individual are the cause of consistent behavior regardless of the situation, the content of test items or interview questions is unimportant (Goldfried & Linehan, 1977). For example, the intent of projective techniques is disguised so that persons will reveal what is troubling them. Behavioral assessment, on the other hand, views fears, irrational thoughts, behavioral excesses or deficits, and attitudes as the very objects of measurement and program planning.

Behavioral assessment meets the requirements of PL 94-142 to a greater extent than does traditional assessment. The development of an IEP requires the specific measurement of current levels of performance in all areas of concern. A thorough behavioral assessment accomplishes this requirement by assessing directly the problems identified by teachers, parents, or students themselves.

An IEP requires statements on the specific objectives to work toward, including behaviors that must be increased as well as decreased. Behavioral assessment requires an assessment of both the problem behaviors and the more adaptive behaviors which the psychologist hopes will emerge.

An IEP requires a detailed description of the nature of services that can help the student to achieve program objectives. Behavioral assessment studies the child's environment to determine which factors are leading to and maintaining problems, and which variables are missing that promote progress.

An IEP must specify how progress toward objectives will be measured. The very techniques of behavioral assessment provide the basis for program evaluation. The following subsection describes how such assessment is carried out.

### Behavioral Assessment of Minority Students

Behavioral assessment appears to meet the needs of educators better than do traditional assessment modes but the question of its nonbiased characteristics remains to be discussed. In accordance with Reschly's conceptualization of assessment bias (see section of module on the nonbiased assessment of mentally retarded students), I believe that behavioral assessment potentially is the least biased assessment approach which can be used to evaluate emotionally disturbed students.

---

### *Task 4*

*Read the following conceptualization of assessment bias. In your own words, describe the major point, which is how assessment bias should be viewed.*

> *A better perspective on nonbiased assessment can be developed by considering the entire purpose of our professional roles. Simple and narrow-minded avoidance of "placement bias" in special education is of dubious value to the minority students who are failing in regular*

*classrooms. The broader issue is providing more effective educational interventions for minority children and youth. Stubborn defense of the current pattern of overrepresentation or strict avoidance of so called "placement bias" are both "nonsolutions" to the challenge of nonbiased assessment. Against this background, the following definition of bias in assessment was developed:*

> *Assessment which does not result in effective interventions should be regarded as useless, and biased or unfair as well if ethnic or racial minorities are differentially exposed to ineffective programs as a result of assessment activities.* (Reschly, 1979)

*The two essential components of this definition of test bias are usefulness and fairness. The first—usefulness—is seen in assessments resulting in effective interventions that improve skills and competencies, and thereby enhance opportunities. This goal should continue to be paramount for school psychologists and special education services. The usefulness of all assessment procedures (including intelligence tests) should be determined on the basis of their contributions to this goal. It is acknowledged, in some instances, that assessments lead to accurate diagnoses for which no known effective interventions can be found. These diagnoses still may be valid (in the sense of validity used by Cromwell, Blashfield, and Strauss, 1975) if they improve estimation of the prognosis or contribute to the prevention of the condition in future cases. However, if effective interventions cannot be developed, accurate prognostic estimates or the prevention of the condition in the future rarely are of benefit to the individual being assessed.*

*Fairness if related closely to the notion of usefulness. Assessments and accompanying diagnoses are interpreted as biased or unfair if they result in the overrepresentation of minorities in programs that are ineffective, or if no interventions at all are planned. Under such circumstances, the diagnosis may be accurate and the assessment conducted competently, but there is no benefit to the individual. Moreover, if a negative connotation or stigma is associated with a particular diagnosis, and if this diagnosis occurs more often with minority children, then the assessments leading to that diagnosis are regarded as biased or unfair. On the other hand, assessments that lead to the accurate description of current behaviors, and to effective interventions should be regarded as fair or unbiased regardless of the racial-ethnic composition of the student groups. Over- or underrepresentation of minorities in various classificattions or programs is not sufficient to establish bias in this conception.*

Reschly's major point is that an assessment process which is directly related to effective intervention strategies has a far greater potential for being nonbiased than an assessment process that leads to ineffective programs or no programs at all. Effective intervention strategies work toward specific, measurable objectives, manipulate immediate environmental events (antecedents and consequents) to help students to achieve desirable goals and include built-in evaluation methods to assess progress. Behavioral assessment, then, potentially minimizes bias because it leads to the development of effective school programs.

Garcia (1975) argued, however, that behavioral assessment and intervention can be used in a discriminatory manner. He suggested that minority students may exhibit patterns of behavior which, although appropriate in their cultural context, may become the targets of behavior modification in the schools. Behavioral techniques are viewed as a means of control through suppression of behavior which is considered inappropriate in the school setting; hence, they can be biased against minority groups. To prevent this, Garcia suggests that a minority representative should be present when any behavior programs are to be planned for minority students. These important points merit further discussion.

The belief that users of behavioral assessment and intervention techniques are solely concerned with the suppression of inappropriate behaviors—the so-called "be still, be quiet, be docile" goals of behavior modification (Winett & Winkler, 1972)—is a misconception. In fact, texts discussing behavioral assessment and therapy clearly indicate that interventions cannot be successful if they seek only to eliminate problems; behavioral assessment and intervention also must be concerned with instilling appropriate coping mechanisms (Ciminero, et. al., 1977). Even a cursory glance at the behavioral assessment literature convinces one that this approach does not preclude the selection of humanistic objectives (e.g., to increase positive thoughts and attitudes and to develop coping behaviors).

In general, there are two types of behavioral change which should be encouraged in the school setting: the reduction of behavioral excesses and the maximization of behavior that is deficient. Excesses usually take the form of behaviors, thoughts or affective responses that clearly interfere with a student's own learning and development and that of others around him/her. Deficits usually occur in coping skills, rational thoughts, and feelings which when learned, promote adaptation and adjustment to school *and* community norms. When school psychologists and educators view target behaviors this way, the minor differences in behavior which may be due to cultural norms will be respected. To the extent that school decision makers are aware of cultural norms, so the selection of objectives for behavioral assessment and intervention can be fairer to minority children.

On the other hand, educators should not make the mistake of neglecting to try to change certain maladaptive behaviors because they are erroneously believed to be culturally-determined. The belief that minority students show distinct culturally determined patterns of behavior in school is open to question. Henderson's section of this module discusses sociocultural considerations in assessment and cautions psychologists against the blind acceptance of the notion of culture specific behavior patterns. Henderson's research suggests that what most educators view as cultural patterns of behavior are, in reality, ethnic stereotypes.

> Although an understanding of the cultural background of children with whom we work is important, blanket descriptions sometimes are more harmful than helpful because the social science research base for the cultural description of ethnic groups is suspect in several aspects. . . . Many descriptions start with the acceptance of dated assumptions without the benefit of first-hand study of the communities in question. (Henderson, Nonbiased Assessment: Sociocultural Considerations)

As Garcia suggested, including parents and other minority representatives if necessary in the data gathering and decision making process will help ensure that parental values are respected and that decisions about behaviors in need of change do not reflect cultural bias.

In sum, several points should be considered in initiating behavioral assessments.

1. The purpose of behavioral assessment of emotionally disturbed students is to provide data that will assist educators to make informed decisions about the goals, objectives, intervention strategies and evaluation techniques of the individualized programs.

2. Behavioral assessment, in contrast to traditional modes of assessment, is suited to meeting the IEP requirements of PL 94-142.

3. Behavioral assessment has the potential to minimize assessment bias when it is used to assess emotional disturbance in minority children.

4. The selection of targets for assessment and intervention is the responsibility of psychologists, educators, and parents who are knowledgeable about school and community norms of behavior. A behavioral assessment requires the selection of objectives to accelerate adaptive and decelerate maladaptive behaviors.

## Assessment of Emotional Disturbance: Specific Guidelines

### Data for Decision Making

Assessments for special education, as with all assessments, are designed and carried out to provide information that will aid in decision making. Thus, before deciding how to assess emotional disturbance we first must identify the decisions that must be made about the student. When the critical decisions are identified, the appropriate assessment strategies to provide the data for these decisions can be selected.

*Task 5*

*Obtain a copy of your state definition and criteria for emotional disturbance.*

*List the decisions that have to be made about a student who is being evaluated for placement in a program for the emotionally disturbed. Next to each decision, describe the data you need to help to make that decision.*

In general, three decisions must be made about children who are referred for placement in a school program for the emotionally disturbed:

1. Prior to placement: Does a child exhibit the behaviors required to classify him/her as emotionally disturbed under state guidelines?
2. Upon placement: What strategies appear to be most appropriate for helping the student to achieve the goals and objectives of the E.D. program?
3. Subsequent determination: Has the child benefited from this program?

The following sections examine each decision separately and describe the techniques for data collection to facilitate making the decisions.

### Classification Decisions

The decision to classify and label a student emotionally disturbed should be based on data supporting an affirmative answer to the following three questions:

1. Do behavioral, cognitive, or affective problems interfere with the student's learning and development and those of his/her peers?
2. Does the student appear to be significantly impaired as a result of these problems in comparison to his/her peers or classmates?
3. Are the kinds of tasks, degrees of classroom structure, and motivational levels that must be provided to help the student beyond the capacity of a regular classroom program?

special educators usually ask school psychologists to provide the data to answer the first question. The data are necessary but not sufficient for classifying a student as E.D. Answers to the last two questions are needed, yet they tend to be ignored in making decisions about labeling and classification.

*Task 6*

*List reasons why you would want data on the last two questions. If you believe they are unimportant, state your reasons.*

Teachers often refer children for behavioral problems and possible placement in E.D. programs when the children's behaviors are like those of other students who are not referred. Why this selective referral occurs is unclear. Caplan (1971) suggest several possible reasons.

Behavioral excesses (e.g., fighting, breaking rules, swearing) interfere with learning and development. However, when such behavior is the norm in a particular peer group, one should be reluctant to classify any child as emotionally disturbed. This caution points up the need for norm-referenced behavioral observation data (Walker & Hops, 1976). The same need holds for behavioral deficits (e.g., withdrawal, shyness, lack of social skills, or lack of interest in school). The question, then, is how unusual or atypical should a student's behavior be, compared to his/her peer group or classmates, before referral should be made?

Concerning question three, there are students whose problems, once assessed, allow affirmative answers to questions one and two. However, a simple intervention (e.g., an individualized home-school report card of behavior or token reinforcement program run by a teacher) can bring about dramatic improvements in school behavior. When the improvement happens, there is little reason to label the child emotionally disturbed and place him/her in a special program. Consequently, before labeling a child E.D., teachers should try certain types of interventions which have proven to be effective and which are feasible, given other demands on the teachers' time. Following a genuine effort at classroom intervention, school psychologists can be more confident in judging whether a student is emotionally disturbed than if no interventions were tried and no data were gathered. Teachers must present evidence that educational efforts to resolve the problems have failed (e.g., a well-thought-out and designed plan that has been tried for a sufficient amount of time, say a week or two, and during which data on behavior were collected).

*Task 7*

*Divide a page into two columns. In the left-hand column list the three questions which are used to decide whether a child qualifies for the E.D. label. In the right-hand column identify the assessment tools that may provide data to answer these questions.*

The following information is needed in order to answer question one:

1. Precise statements about the child's behavioral, cognitive or affective problems (including academic skill deficits).
2. Data on the magnitude of these problems in terms of frequency, rate, duration, intensity, etc.
3. Data on how the student's problems affect his/her peers.

The following data are required to answer question 2:

The extent of similar problems among peers or some appropriate normative group.

The following data are required to answer question 3:

1. Tasks the student attends to and succeeds at.
2. The degree of environmental structure in which the student appears to work best.

3. Rewards and punishments that best motivate the child.
4. The extent of previous educational efforts to help the child.

Data to answer the first question can be gathered from interviews with teachers, parents, peers, and the student; direct observations in the classroom; behavior checklists and other self-report measures (Cautela, 1977); and norm- and criterion-referenced achievement test data.

Data to answer the second question can be gained from interviews, direct observation, and behavior checklists.

Data to answer question three are best obtained by observing the student in settings that differ in structure, tasks demanded, and motivators used. Simulated settings or situations can be used (Nelson, 1977). To evaluate the results of past educational interventions, examine data on the student while the program was in effect and compare to baseline measures; also, examine the information suplied by the teacher or school psychologist on the extent to which the program was implemented.

The following report of an actual behavioral assessment for emotional disturbance demonstrates how the techniques can be used to answer the three basic questions and decide if a student qualifies for an E.D. label.

## Behavioral Assessment

Name: Terry Stones
D.O.B.: 2/1/68
Dates of assessment: 10/5-10/19

**Problem Identification:**

Terry was referred by his classroom teacher, Mrs. Cash, because of extensive acting-out behavior in class. In the 30 minute interview with her she specified acting-out behavior as getting out of his seat at least 30-40 times a day, hitting, touching other students, calling out, using abusive language, and refusing to follow directions. She also stated that Terry has no friends. The other students dislike him because he is constantly bothering them and getting them into trouble.

Mrs. Cash stated that she has several other children who break classroom rules, but none to the extent that Terry does. She reported that Terry has exhibited these behaviors since the first week of school and that his cumulative folder details a similar history since second grade.

Terry's reading and math skills are more than two years below grade level. According to Mrs. Cash, he appears totally unable to converse in any pleasant manner with adults or peers. Based on my interview with Mrs. Cash, the following problems were identified:

1. Behavioral excesses: fighting, out-of-seat, cursing, general disruptiveness.
2. Behavioral deficits: no skills in making friends or communicating with adults.
3. Academic deficits: below grade level in reading and math.

Mrs. Cash filled out the Peterson-Quay behavior problem checklist, the results of which were in agreement with her descriptions of Terry's behavior.

**Behavioral Observations:**

Mrs. Cash and I observed and recorded Terry's classroom behavior for one week. Mrs. Cash kept a frequency count of disruptive behavior for 5 days. Terry averaged 50 disruptive acts a day. I observed Terry for 6 randomly chosen intervals of 10 minutes duration over a period of 2 days. My observations validated those of Mrs. Cash. Using a 20-second interval recording pace, Terry was observed to engage in disruptive behavior in 25% of the intervals. In addition, I recorded disruptive acts by other classmates. The total amount of disruption from others amounted to only 5% of recorded intervals.

Other school personnel were asked to observe and record any instances of the following behaviors: conversing with adults, cooperative play with other students, initiating social interactions with peers or adults. They observed for one week in the cafeteria, play areas, and during recess. Terry was never observed conversing with adults, playing cooperatively with other students for more than 10% of the time, or initiating any social interactions.

**Interview with Terry:** (30 minutes)

This interview took place in the counselor's office. I explained to Terry who I was, why I was interviewing him, and that I had observed him in class. Terry was aware that he was "in trouble" but he believed that he was being picked on by his teachers. He was tired of being given "baby readers to read and kindergarten math." When I explained that his reading and math skills were not up to sixth grade level, he replied that other kids made fun of him because he was doing what the little kids do. He said he just can't learn to do reading and math well. He reported that he can't wait to get out of the "stupid school" and get to junior high where things were nicer. At junior high, he said, you travel to different classes and don't sit bored with the same "old bitch" all day.

He believed that the principal, Mr. Bran, picked only on the black kids. He also indicated that his parents are fed up with the constant harassment by the school in the form of notes sent home and phone calls. When I told him that a lot of what I observed in class was started by him and I saw no one picking on him, he retorted, "Well, then, you need new glasses."

I asked if he felt there was any way the situation could be improved. He said, "Yes! When I get out of this damn school." When I explained to Terry that he had 8 months of school left and that the situation can't remain as it is, he indicated that "it's up to them (teachers and principal) to make it better." He expressed little sense of control over the school situation.

From this interview, the following possible problems can be identified:

1. Irrational thoughts that he is picked on because of his race, that teachers have it in for him.
2. Self-defeating self-verbalizations: i.e., can't do reading and math, can't control the situation.
3. Affective responses: "Others are laughing at me and that's awful"; concerns over what others think of him.

**Assessment of environment-behavior interaction:**

Terry's behavior varies throughout the school day, apparently depending on such variables as difficulty of tasks assigned, reinforcers in the environment, classroom structure in terms of specificity of rules and consequences for breaking them, and amount of adult attention. From observations and teacher reports, it appears that Terry's behavior is worse in settings where adult attention to behavior is sporadic and tasks are inappropriate for his ability level. Mrs. Cash has 31 students in her class and has difficulty attending to Terry when he is behaving well and ignoring many of his attention-getting behaviors. Likewise, Terry's peers are a constant source of reinforcement for inappropriate behavior. The only consequences the school uses for bad behavior are "timeout" in the principal's office and a letter home or parent conference. None of these work with Terry because he wants to get out of class and his parents are at odds with the school.

PE class is another story. Terry enjoys the activities, receives much praise and encouragement from the teacher, and the class has explicit rules and routines which Terry respects. The consequences for misbehavior also are explicit: removal from the gym.

The remedial reading teacher reports that Terry is no problem in her classes which run for 30 minutes. Terry works in a group of six, in a reader at his level. His teacher also has set up a token reinforcement program to which

Terry responds very well. She is quick to praise him for appropriate behavior and knows which behaviors to ignore and which to punish by loss of tokens. Terry has averaged over 90% of his tokens for the past 4 weeks.

**Results of trial effort to improve Terry's behavior in Mrs. Cash's room:**

Despite some reluctance, Mrs. Cash agreed to the use of a good behavior checklist for seven school days to see if Terry's behavior would improve over what we observed at baseline. Four simple rules were specified, and Terry received a check mark every 15 minutes for the absence of any rule violations. Mrs. Cash also was to make an effort to give him more social praise. 80% of all possible checks allowed Terry extra gym time every day. Terry received 80% the first day, 75% the second, 50% the third and fourth, and then refused to keep the checklist anymore. His inappropriate behavior returned to the level it was at baseline.

Failure may be attributed to a number of factors: Terry's academic tasks still were too difficult for him; peer attention for inappropriate behavior continued; Mrs. Cash was too busy with other students to regularly give Terry social praise; Terry's beliefs that Mrs. Cash doesn't like him remained unchanged. The program could be improved, but it would require more time and attention than Mrs. Cash could give.

**Intellectual skills assessment:**

During both IQ and achievement testing, Terry behaved in a manner similar to that described by his remedial reading teacher: He attended to directions and tasks even though some were difficult and he expressed pleasure at doing well. On the Brigance Skills Inventory he worked to his frustration point but expressed such comments as, "I can't do well on these." These comments were particularly evident on math problems. Terry clearly demonstrated that he can control his behavior in situations requiring sustained attention to sometimes tedious tasks, given close adult supervision, liberal social praise, and attention for appropriate behavior.

Terry's IQ on the WISC-R was well within the average range, indicating ability to succeed at academic tasks.

Results of assessment using the Brigance indicated skill levels in reading at late second grade and in math at early third grade. A more detailed breakdown of the specific skills that Terry has and has not mastered can be ascertained from the attached record forms.

**Summary of problem areas in need of change:**

**Behavior:** Terry must learn to:

1. stay in seat and complete work,
2. speak out at appropriate times,
3. initiate and engage in conversations with adults,
4. initiate and engage in appropriate social interactions with peers,
5. ask for help when he cannot do work, and
6. improve in reading and math skills.

**Affect:**

1. improve attitude toward school, and
2. express anger in socially appropriate ways.

**Thoughts:**

1. decrease irrational thoughts (e.g., "Teachers don't like me. The principal picks on me because I am black. Everybody picks on me."),

   decrease self-defeating beliefs (e.g., "I can't learn to read well."),
3. increase positive self-statements related to academic skills, and
4. develop more realistic beliefs about why he is failing in school and why he is being disciplined.

**Interpersonal relationships:**

1. increase friendships, and
2. play games cooperatively.

**Recommendations:**

Terry's school-related problems are significant enough to impede his learning and the learning of others in his classroom. His behavior is markedly worse than that of others in his classroom. This is particularly evident in large-group situations where adult attention is sparse. The learning tasks that Terry can succeed at, the level of adult attention required to keep him on task and learning, and the immediacy of reinforcement and punishment needed are beyond the resources of a regular classroom setting. He also requires cognitive therapy to change certain thoughts and beliefs that interfere with his adjustment to school rules and school learning. For the above reasons the following recommendations are offered:

1. Terry should be placed in a special education program under the classification of emotionally disturbed, pending parent approval.
2. The objectives of such a program should include those listed under "Summary of problem areas in need of change."
3. The details of this program should be worked out in a joint meeting with teachers, parents, school psychologist, and special education staff.
4. The program should include sufficient mainstreaming to allow the evaluation of progress toward program objectives using data-gathering techniques described in this report.
5. [1]At the meeting described in #3, strategies to help Terry to meet objectives and methods to evaluate progress will be mutually discussed and decided upon.

---

[1] As a general rule, specific details of a program should be worked out jointly with those responsible for its execution. This helps to dispel the all-too-prevalent notion that school psychologists are unrealistic and dogmatic about their recommendations and insensitive to the demands of classroom teachers.

---

### *Task 8*

*Reread the report of the more traditional assessment in the first section of this paper. Compare the two reports with respect to their presentation of data to help to decide if the students qualified for an E.D. label. What shortcomings do you see in the behavioral assessment?*

---

When we compare the two reports, striking differences are noted. Behavioral assessment makes less use of standardized test data and no use of projective tests. Rather, it relies on observations by teachers and psychologists, interviews with the student and his teachers, and a trial intervention program. The problems that initiated the referral are the foci of the assessment and intervention. An assessment of how the immediate environment appears to affect behavior is included. Although more time appears to be taken by the behavioral assessment, this is not the case. The actual time devoted to assessment by the psychologist was five hours: 90 minutes in interviews with the regular classroom teacher, the child, and the PE and reading teachers; 45 minutes in consultation with Mrs. Cash to set up the trial intervention; 60 minutes to give an IQ test (the skills inventory was administered by an aide), and the rest to do observation. In the behavioral assessment school personnel were relied upon far more for the data collection.

The behavioral assessment could have been improved by the following additions:

1. Systematically recording Terry's behavior in PE and remedial reading class.
2. Systematically recording social skills on playground, during recess, and in cafeteria.
3. Observations by neutral observers.
4. Arranging to observe Terry in standardized, simulated settings that used different environmental structures, tasks, and motivations.
5. Using a self-report instrument to assess better affective and cognitive dimensions.
6. Including parents more directly in the diagnostic-intervention process.

The judgment of the psychologist, however, was that the data were sufficient to decide that Terry qualified for the ED label and an ED program. Of course, no classification decision can be made without the agreement of the other members of the child study team.

Time and case load requirements always will force school psychologists to seek a compromise between what is optimal and what is feasible in assessment. In any case, the bottom line is that a judgment must be made, and we hope that the judgment will be based on defensible data that allow one to answer the questions at the beginning of this subsection.

### Programming Decisions

The preceding section provides examples of collecting data to facilitate decisions on labeling. Unfortunately, many assessments of emotionally disturbed students provide data to help to classify the students but they give educators and psychologists few clues for how to deal with the children. The assessment process, however, is helpful to students and their teachers to the extent that it provides data useful to program planning.

The psychologist's assessment should include information on those educational tasks in which the student experiences success; the degree of classroom structure needed to complete these tasks; and the motivational strategies necessary to help the student initiate, persist at, and complete them. In other words, a functional analysis of the antecedent and consequent events that influence the student's learning goes hand in hand with an assessment of the problem area.

A functional analysis can be made in several ways. Typically, it is carried out through a systematic interview with a teacher and systematic observations of the student in a variety of settings. Ideally, the analysis derived from the teacher interview would be followed by observations in classrooms and those settings where the problem behaviors occur. In reality, heavy reliance must be placed on the teacher's verbal reports, supplemented by the psychologist's brief observations and the data records kept by teachers. In some cases, however, even these latter sources of data cannot be obtained.

Inasmuch as the interview may be the primary source of data for a functional analysis, a detailed example of an interview follows; comments are included on how to elicit the essential information and analyze it in functional terms. This procedure often is called a "problem identification interview" (Bergan, 1977).

## Problem Identification Interview

| Interview dialogue | Thoughts during the interview |
|---|---|
| *Interview Objective 1:* **Elicit behavioral descriptions of problems.** | |
| Psychologist: I see from your referral that Terry has been a problem for some time. Tell me about the kinds of things he's been doing in school that have you all so concerned. | Let's get focused on school related issues rather than his home life, etc. |
| Teacher: Terry's been a problem ever since the start of school. There are a lot of problems at home and I don't see what we can do for that boy as long as his parents don't care and don't discipline him. | Ignore the comments about the home. Give her a firmer cue to discuss school-related problems. |
| **P.** Well, what are some of the behaviors that are disrupting your class? | |
| **T.** He's never in his seat, he's always pushing and touching the kids around him, his language is atrocious, he calls out when he wants to—just general classroom disruption. | |
| **P.** Certainly not one of your top candidates for the good conduct award! What other things about his behavior have you observed that concern you? | |
| **T.** He never does any school work and I never see him with any friends. During recess or on the playground or in the cafeteria he's always by himself. I really think he lacks confidence and feels terrible about himself. He'll never talk to adults either. He just kind of puts his head down and has this sullen expression whenever you try to talk to him about his behavior. | Good. Now summarize the problem list. Also check if he has friends at home. He might lack social skills or just dislike school so much that he makes no friends. Or maybe she's never observed him when he's around other children or adults. Better question a little further. |
| **P.** So he's never in his seat, he's always pushing and touching other kids, his language is offensive, calls out whenever he wants to, doesn't do his work, and is by himself all the time. Are there other concerns? | |
| **T.** That seems to be it. | |

| | |
|---|---|
| **P.** You mentioned that you never see him with other children. Do his other teachers report the same thing? | |
| **T.** I spoke with the PE teacher and some of the monitors in the lunchroom and they say he's always by himself. On the playground, whenever he's with other kids, it's in a fight or some kind of argument. | Better check with them and have them observe more closely. |
| **P.** Do his parents indicate that he has friends at home? | |
| **T.** I never talk to his parents. That's the basis of the whole problem—they just don't care. Why, you wouldn't believe the trouble I've gone through to work with them. | You blew it! Better get the interview back on track. |
| **P.** How are his academic skills? | |
| **T.** What academic skills? He has no reading or math skills. He's in remedial reading but is getting no help in math. I've got too many kids to give him the one-to-one that he needs. You can check with the reading teacher on his reading skills. She also gave him some IRIs which will give you some idea of how low he is. | |
| **P.** I'll do that. So there is definitely a problem with his skill level. It's not just that he refuses to do work. Would you say there's both a motivational and a skills problem? | |
| **T.** Definitely. Some of the work he just can't do. Some he can, but won't do it. | I wonder if he is the only student in her class with such problems or if there are others like him and she's singling him out for some reason? |
| **P.** Would you say that Terry's classroom problems are noticeably worse than other students in your class? | |
| **T.** Definitely. My class is pretty well behaved for the most part. Nothing perfect, but it's tolerable. I'm not scolding kids all day. | She seems to be willing to tolerate some disruption. So her standards for good behavior seem realistic. Classroom observations will validate these impressions. |
| *Interview Objective 2:* **Elicit behavioral description of goals.** | |
| **P.** What kind of behavior do you expect from Terry, Mrs. Cash? | |
| **T.** I want him to follow rules just like anybody else. | She must think that's a pretty dumb question. Better rephrase it. |
| **P.** What are the rules of your class and when do you go over them? | |
| **T.** The students are supposed to come to class with all material required to do their work. They should stay in their seats unless given permission. Raise their hands to speak. Complete all work before the end of the period. | Good! Praise her for a clear set of rules! |
| **P.** Those are clear and well stated in a positive manner. And those are the rules you want Terry to follow? Are there other goals you have for Terry? | |
| **T.** Yes. I'd like to see him initiate conversation with adults and play nicely with other students. I think he'd enjoy school more if he did. | O.K. I seem to have a clear idea of the problems and expectations. Let's find out how conducive the environment is to promoting good behavior. |
| *Interview Objective 3:* **Elicit antecedent events that may signal inappropriate behavior.** | |
| **P.** I'm interested in knowing if things happen in class that set off the disruptive behavior. What are you doing, for example, just before he calls out or gets out of his seat? | |
| **T.** Well—if I'm near him or explaining things to him, he's O.K. But just let me turn my back or write on the blackboard and then something will happen. Also, when I hand out a math or reading worksheet, he'll usually yell out something like, "I don't understand this," or "This is too hard," or "Same old crap." | Proximity seems to be an important condition. Also, he may not be able to do the worksheets. I wonder if his peers instigate his behavior in any way? |

| | |
|---|---|
| **P.** When he pushes or touches other children, is there anything that might set that off? | |
| **T.** Not as far as I can tell. The other kids don't ever start anything with him. He can't stand correction. Whenever I correct him for anything or tell him to do something, he'll usually make some snide comment. | Better summarize antecedent conditions. |
| **P.** So his behavior gets worse when your back is turned, when you're far away from him or when you have to correct him or tell him what to do. Also, getting work he can't do seems to set off his calling out. Is that correct? | Implication is that closer adult supervision and work at his level might improve his classroom behavior. Better check that out with other teachers. |
| **T.** Yes. | |
| *Interview Objective 4:* **Elicit specification of consequent conditions that may be maintaining the inappropriate behavior.** | |
| **P.** What do you do when he gets out of his seat? | |
| **T.** I tell him to sit down. | |
| **P.** And if he refuses? | |
| **T.** Well, sometimes I'll give him a warning, sometimes I'll ignore him for a while, sometimes I escort him back. When things really get bad, I'll send him out of the room to the principal's office. | Seems like Terry is pretty much in control. If he pushes her far enough he gets just what he wants—out of a class he hates. She has no consistent and effective way of responding. There may be no consequence he cares to avoid. |
| **P.** From your reports, though, I take it that these things haven't cut down on his disruption. How about when he calls out? What do you do? | |
| **T.** I ignore it as long as I can. Then I go over and speak to him. Again, if it's really disrupting the class, I send him from the room. | |
| **P.** What do other children do when Terry disrupts the class? | |
| **T.** They'll tell him to be quiet, or sometimes make fun of him when he's sent from the room. | Peers also are reinforcing his behavior. It's going to be hard to improve this situation in the present classroom environment. |
| **P.** There seem to be no consequences either for misbehavior or failure to do work that Terry wants to avoid? | |
| **T.** Yes. And when I get no home support, what can I do? | I wonder if this kid gets any positive strokes at all in this class? |
| *Interview Objective 5:* **Elicit antecedent and consequent conditions surrounding appropriate behavior.** | |
| **P.** Are there any times when Terry is quiet, doing his work, following the rules? | |
| **T.** Yes, but they are few and far between. | |
| **P.** For example, if Terry gets an assignment from you and starts to work at it, what do you do? | |
| **T.** Well, if he completes it correctly, I praise him for it. But I just don't have time to be right next to him encouraging him and explaining things. The other kids work independently. Why can't he? | Maybe if he had more attention for good behavior, he might stay on task longer. But I see her point. Can she attend to him more given 30 other students with needs? |
| **P.** What kinds of tasks is he likely to work on? | |
| **T.** Things that he can do . . . which isn't much. | Well, the situation isn't hopeless; he'll work at things he doesn't become frustrated at immediately. |
| **P.** Any particular time of day or week when he works best? | |
| **T.** Well, if he works at all, it's after PE and remedial reading. Mondays and Fridays are his worst days. | Maybe he comes out of these classes feeling some sense of competence and saying positive things to himself. Maybe Mondays and Fridays are days during which greater structure is needed. |
| **P.** Have you ever spoken to him in any way that seems to improve his behavior for a day or two? | |

**T.** Well, there have been times when I've talked to Terry before the class and told him explicitly what I expect from him, and that if he can behave today, I'll let him be line monitor or messenger. That keeps him going, but it's only temporary.

So the old Premack Principle works sometimes. I wonder what reinforcers work for him.

**P.** What rewards work best for him?

**T.** Nothing for any length of time. He likes to do things like be monitor, or erase boards, or be messenger. But sometimes those mean nothing to him. He'd probably spend the whole day in gym if he could.

Good information! Better have Terry fill out a reinforcement menu before we plan any program. At least he responds to activity reinforcers and isn't at the primary level. Better check with PE and reading to see how they motivate him.

**P.** So there are rewards that work with Terry and there are times he'll behave tolerably. The problem is how to get more consistent good behavior from him.

**T.** I guess that about sums it up.

**P.** Thanks for your time, Mrs. Cash. I'll get back to you as soon as I've consulted with Terry's other teachers.

The interview serves three main purposes: (a) It allows the psychologist to identify certain targets for behavior change. (b) It gives the psychologist some understanding of how environmental conditions affect Terry's behavior and allows him/her to generate numerous hypotheses about how to decrease certain behaviors and increase others. (c) It pinpoints areas in need of further observation and helps to focus the interviews with other school personnel.

Specifically, the following information has been learned from the interview to help in program planning:

1. Attention from adults and peers is important to Terry. Attention now reinforces inappropriate behavior. In a well-designed ED program, social reinforcers could be used to strengthen appropriate behavior.

2. Terry will work at tasks at which he can succeed. This suggests the need for thorough skills assessment that pinpoints the exact levels at which Terry can experience success. Also, in line with recent research, improving academic skills may be the best way to eliminate certain behavioral excesses.

3. There are certain activity reinforcers that will motivate Terry to behave well. He does not have to be dealt with at a strictly primary reinforcement level.

4. Terry responds to structure and close adult supervision. His behavior deteriorates when teacher attention is sparse following good behavior and during unsupervised periods.

Interviews with PE and remedial reading teacher suggest the following:

5. Terry does not require strictly individualized attention. He will work in small groups provided that the work is at his success level and that rewards and punishments are clear and immediate.

6. He will behave in large groups if the activities are those he enjoys and if the consequences for misbehavior are immediate removal from these activities.

Classroom observations suggest the following:

7. Terry's behavior is being reinforced by teacher and peer attention. There are not meaningful consequences for misbehavior. The regular classroom cannot provide the amount of help he needs to master academic skills nor the level of structure and motivation he needs to learn appropriate behaviors.

The preceding information can be related directly to program development. It clearly illustrates the close link between behavioral assessment and educational programming.

### Collecting Evaluation Data

Assessment procedures, particularly those that determine pre-intervention levels of behavioral excesses and deficits, are not put aside when a student is placed in an ED program. Rather, they continue to be used to monitor objective attainment. Thus, the assessment procedures which are used to help to determine that a student is qualified for an ED program also are used to determine when the student no longer needs the program.

Data collected through classroom observation, behavioral checklists, interviews, and skills assessment are pre-program baseline measures, the starting point to gauge progress toward the objectives of the ED program.

For this evaluation to be valid, the assessment items and the conditions of assessment at pre- and post-evaluation dates must be essentially similar. This requirement presents obstacles to self-contained ED programs that do not provide for mainstreaming. ED programs using the resource room concept are easier to evaluate. In such programs, the student's behavior when he is in regular classes can be compared to pre-program levels. Pre-post changes should be sought on classroom observation instruments, behavior checklists, and skills inventories. Measuring changes from interviews presents problems; however, teacher interviews can be content analyzed for changes in reported problem behaviors as well as changes toward more appropriate behaviors. Similarly, the student's verbalizations during interviews can be content analyzed for changes in irrational cognitions, negative self-statements, and attitudes toward teachers and school. Psychologists evaluating progress should be concerned with convergent validity when they examine the results of all the assessment techniques.

### Remaining Questions About Behavioral Assessment

A major goal of this paper is to encourage the use of behavioral assessment as an alternative to the traditional assessment of emotional disturbance. Some questions about behavioral assessment which remain to be answered, are discussed in this subsection and the answers, I hope, are satisfactory.

*Should you confine your assessments to the specific observable behaviors that lead to the initial referral?*

Teachers generally refer a student for a possible diagnosis of emotional disturbance because they have observed certain behavioral excesses or deficits. Such observable events are necessary but not sufficient data upon which to base decisions of classification and program planning. A comprehensive assessment should include both overt and covert behavioral events.

*How does the use of self-reports in behavioral assessment differ from the use of self-reports in more traditional assessments?*

Traditional approaches to assessment focus on personality descriptions and attempt to assign children or adolescents to particular diagnostic categories. The self-reports of feelings, impressions, fears, attitudes, or beliefs are of interest to the extent that they give psychological clues to certain underlying personality constructs that lead to a diagnosis of mental health or illness.

In contrast, the school psychologist can and should interpret self-reports as verbal behaviors of intrinsic interest which can become the focus of classroom interventions. "When self-reports are verbal behaviors that are themselves the main focus of interest, the question of their validity as indicators of other phenomena is not central" (Cautela & Upper, 1976). Thus, a student's reports of fears, beliefs, attitudes, sensations, images, or interpersonal behavior patterns themselves provide some of the very criterion variables that form goals and objectives of assessment and intervention.

*What are some systems of available assessment that assure thorough coverage of all relevant areas of student behavior?*

Over the last decade, several comprehensive, multi-dimensional assessment systems have been developed that are relevant to the assessment of emotional disturbance in schools (Goldfried & Pomeranz, 1968; Kanfer & Saslov, 1969; Lazarus, 1973; Wolpe, 1969). All these systems have in common a concern for overt behavior and for cognitive and affective states. Lazarus's system—BASIC ID—can be adapted very easily by school psychologists who carry out ED assessments. Lazurus recommended that psychologists systematically gather data on the dimensions listed in Table 1 using interviews, self-reports, and behavioral observations.

Although many of the self-reports cited in Table 1 were designed for use by clinical psychologists with adult patients, they easily can be adapted for school use.

*How valid and reliable are the data generated through systematic observation, interviews, or self-reports?*

Data from observations are valid if they represent the "true behavior" of a student when he is not being observed. Observational data have been subject to the same reliability procedures and requirements as have paper and pencil tests. Stability, inter-observer agreement, and internal consistency data have been derived for a variety of observational procedures (for a review of studies that report this research, see Haynes, 1978). In summing up these studies, one can conclude that although numerous sources of error can affect the reliability of observational data, these sources can be controlled and estimated. Consequently, if proper procedures are followed, school psychologists can have confidence in the reliability of their school observations (Doke, 1976; Mann, 1976).

Adequate reliability is a necessary but insufficient condition for judging the adequacy of observational data. In addition, one should be concerned with content and criterion-related validity. Content validity indicates how well your observations sample the behaviors of interest. Criterion-related validity indicates how well instruments that measure similar behaviors correlate with observations. Several studies suggest that systematic observational systems measure the behaviors they are designed to measure.

Interviews are valid and reliable if the teacher's or child's report of behavior gives an accurate picture of behavior in the natural setting. The conclusion reached, after reviewing what scant research exists, is that interviews provide data of questionable validity. Reliable and valid data can come from an interview, depending on whether the following factors are controlled (Haynes, 1978): interviewer skills, interviewer status, reactivity, topic of discussion, and extent of information on the interviewee. The assumption is that if such sources of bias are controlled, the interview can provide valid data.

The picture acquired of a student from a questionnaire should reflect the student's behavior in the natural environment. When self-reports of fears, anxieties, cognitions, images, or sensations are viewed as behaviors in themselves, self-reports have been shown to be as valid and reliable as other types of measurement (Cautela & Upper, 1976).

## Conclusion

The assessment process should be more than an academic endeavor. It should be the key to unlocking the potential for growth and development in everyone assessed. This paper proposes behavioral assessment as the best designed procedure to help educators plan programs to help students learn and develop while minimizing bias toward minority students. As the saying goes, "Exceptional students need exceptional educational programs." The guidelines presented in this paper, if followed, should allow school psychologists to help to develop such exceptional programs.

## Table 1
## Suggestions for Data Gathering

| Dimension | Assessment Questions | Suggested method of measurement |
|---|---|---|
| 1. Behavior | What behaviors of the child occur at such a high rate that they impair school learning and development? What behaviors are absent from the child's repertoire which, if present, would promote school learning? What behavioral assets does the child demonstrate? | Behavioral interview (Bergan, 1977); Behavior checklists (Novick, et al., 1966; Werry, 1968; Quay, et. al., 1966; Dielman, et. al., 1971; Spivack & Swift, 1966); Behavioral observations (Mann, 1976); Reinforcement survey schedule (Cautela, 1977). |
| 2. Affect | To what extent do fear, anxiety, anger, tension, etc., prevent the child from learning in the regular classroom? | Fear Survey Schedule; Cues for tension and anxiety survey schedule; Behavioral analysis history questionnaire; Depression adjective checklist (Cautela, 1977). |
| 3. Sensations | Does the student have any physical pains, headaches, ailments, etc., that affect school performance? | Behavioral analysis history questionnaire; Behavioral self-rating checklist ( Cautela, 1977). |
| 4. Images | Does the child regularly have certain fearful or upsetting images which interfere with his school learning? | Imagery survey schedule (Cautela, 1977). |
| 5. Cognitions | What beliefs, ideas, insights, etc., does the student have about school, certain teachers, peers, that may interfere (or facilitate) with school progress? | Thought-stopping survey; Behavioral self-rating checklist (Cautela, 1977). |
| 6. Interpersonal relationships | Does the child have friends? What prevents the development of friendships? Does the child interact with peers or adults? What is the nature of this interaction? | Social avoidance and distress scale (Cautela, 1977). |
| 7. Drugs | Is the child in need of certain medications or health regimens that can improve school performance? Is the child taking any medication or drugs that may explain certain behaviors? | Drug Questionnaire (Cautela, 1977). |

## POSTTEST

1. Mental health professionals in general agree on the early signs of ED.
   True
   False
2. With respect to the causes of emotional disturbance, we can conclude,
   a. The weight of research traces the causes to disturbances in family functioning.
   b. Research seems to support the view that the causes of ED are found in the immediate environment of the child.
   c. Conceptualizations of causality of ED can affect teaching behaviors.
   d. Most emotionally disturbed children show signs of minimal brain dysfunction on neurological exams.
3. If objective tests are used, psychological assessment minimizes subjectivity by the psychologist.
   True
   False
4. One's decision to label a child ED will be influenced by
   a. How upset teachers and administrators are at the behavior of the student.
   b. What aspects of the child's behavior adults tend to focus on.
   c. The value system of those judging ED.
   d. Primarily the results of test data.
   e. a, b, c.
5. Behavioral assessments rely almost exclusively on the measurement of operant behavior.
   True
   False
6. The selection of target behaviors that become the foci of behavioral assessment
   a. Usually involves behavioral excesses that must be decreased.
   b. Is dictated by the very nature of behavioral assessment procedures.

c. Must include behaviors to increase and decrease.
d. Often has resulted in the elimination of cultural patterns of minority behavior.

7. According to Reschly, assessments are biased if they lead to ineffective programs for minority children.
True
False

8. Behavioral assessments for emotional disturbance should focus on
a. The specific covert and overt behaviors that initiated the teacher referral.
b. Only those behaviors that can be affected by school-based programs.
c. Behavioral, cognitive, and affective dimensions of student performance.
d. Cognitive and affective events less than overt behavioral ones.

9. Projective tests have low decision-making power for classroom teachers.
True
False.

## SIMULATIONS

1. Imagine that you are responsible for a plan to help Terry to learn appropriate ways to behave in school. On the left-hand side of a paper list the section in the traditional report that helps you to decide how to help him. Under that, describe a teaching strategy derived from that section. On the right-hand side of the page, list sections from the behavioral assessment (including the interview) that help to decide how to plan a program. Under that, specify a teaching strategy derived from that section. What do you conclude about the relevance of traditional assessments vis-a-vis educational programming in comparison to behavioral assessments?

2. Break into groups of three. In each group the members role play a school psychologist, teacher, and neutral observer. Conduct problem identification interviews, modeled after the sample interview in this paper. The role of the observer is to insure that all the objectives of the interview are attained. Each member should play every role. The students referred for emotional disturbance should display a variety of behavioral, cognitive, and affective problems.

3. Conduct a panel discussion with a school psychologist, regular education teacher, special education teacher, principal, and parent on the panel. The audience should pretend to be parents at a PTA meeting who want to question the school's procedures for the assessment of ED. The audience should question the panel on the following issues:
a. Overrepresentation of minorities in the ED program.
b. Validity and reliability of the assessment methods.
c. Assurances of effective ED programs.
d. Information related to student progress in the program.
e. Effects of the ED label on the child once out of the ED program.
f. Who decides on the objectives of the program?

## KEY TO PRETEST

1. c
2. False
3. c
4. True
5. b
6. False
7. e
8. True
9. d
10. False

## KEY TO POSTTEST

1. False
2. c
3. False
4. e
5. False
6. c
7. True
8. c
9. True

## REFERENCES

Bardon, J. I. Overview of issue implications for future trends in school psychology. *Journal of School Psychology*, 1972, *10*, 207-11.

Begelman, D. A. Ethical and legal issues of behavior modification. In M. Hersen & R. M. Eisler (Eds.), *Progress in behavior modification* (Vol. 1). New York: Academic Press, 1975.

Bergan, J. R. *Behavioral consultation*. Columbus, Ohio: Merrill, 1977.

Bergan, J. R., Byrnes, I. M., & Kratochwill, T. R. Effects of behavioral and medical models of consultation on teacher expectancies and instruction of a hypothetical child. *Journal of School Psychology*, 1979, *17*, 307-16.

Caplan, G. *The theory and practice of mental health consultation*. New York: Basic Books, 1971.

Cautela, J. R. *Behavioral analysis forms for clinical research*. Champaign, Illinois: Research Press, 1977.

Cautela, J. R., & Upper, D. The behavioral inventory battery: The use of self-report measures in behavioral analysis and therapy. In M. Hersen & A. Bellack (Eds.), *Behavioral assessment: A practical handbook*. New York: Pergamon, 1976, 77-110.

Ciminero, A. R., Calhoun, K. S., & Adams, H. E. (Eds.). *Handbook of behavioral assessment*. New York: Wiley, 1977.

Dielman, T. E., Cattell, R. B., & Lepper, C. Dimensions of problem behavior in the early grades. *Journal of Clinical and Consulting Psychology*, 1971, *37*, 243-49.

Doke, L. A. Assessment of children's behavior deficits. In M. Hersen & A. Bellack (Eds.), *Behavioral assessment: A practical handbook*. New York: Pergamon, 1976, 493-536.

Fremont, T. S., Klingsporn, M. J., & Wilson, J. H. Identifying emotional disturbance in children—The professionals differ. *Journal of School Psychology*, 1976, *14*, 275-82.

Garcia, D. Behavioral principles with minority children. *The Recorder*, 1975, *4*, 11-15.

Goldfried, M. R., & Linehan, K. S. Basic issues in behavioral assessment. In A. R. Ciminero, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment*. New York: Wiley, 1977.

Goldfried, M. R., & Pomeranz, D. M. Role of assessment in behavior modification. *Psychological Reports*, 1968, *23*, 75-87.

Goodwin, D., & Coates, T. *Helping students help themselves*. Englewood Cliffs: Prentice-Hall, 1976.

Haynes, S. N. *Principles of behavioral assessment*. New York: Gardner, 1978.

Hewett, F. The engineered classroom. Fourth Annual Education Engineering Conference. School of Education, UCLA, 1968.

Kanfer, F., & Saslov, G. Behavioral diagnosis. In C. M. Franks (Ed.), *Behavior therapy: Appraisal and status*. New York: McGraw Hill, 1969.

Lazarus, A. A. Multi-modal behavior therapy: Treating the BASIC ID. *Journal of Nervous and Mental Diseases*, 1973, *156*, 404-11.

Mann, R. A. Assessment of behavioral excesses in children. In M. Hersen & A. Bellack (Eds.), *Behavioral assessment: A practical handbook*. New York: Pergamon, 1976, 459-92.

Novick, J., Rosenfeld, E., Bloch, D. A., & Dawson, D. Ascertaining deviant behavior in children. *Journal of Consulting Psychology*, 1966, *30*, 230-38.

Nelson, R. O. Assessment of child behavior problems. In A. R. Ciminero, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment*. New York: Wiley, 1977.

Quay, H. C., Morse, W. C., & Cutler, R. L. Personality patterns of pupils in special classes for the emotionally disturbed. *Exceptional Children*, 1966, *32*, 297-301.

Spivack, G., & Swift, M. The Devereux elementary school behavior rating scale. *The Journal of Special Education*, 1966, *1*, 71-90.

Tombari, M. L., & Bergan, J. R. Consultant cues and teacher verbalizations, judgments, and expectancies concerning children's adjustment problems. *Journal of School Psychology*, 1978, *16*, 212-19.

Tombari, M. L., & Brein, S. Psychological reports: Do the recommendations fit the data? Paper presented at the convention of the National Association of School Psychologists, New York, 1978.

Walker, H. M., & Hops, H. Use of normative peer data as a standard for evaluating classroom treatment effects. *Journal of Applied Behavior Analysis*, 1976, *9*, 149-68.

Werry, J. S. Observing the classroom behavior of elementary students. *Exceptional Children*, 1968, *35*, 461-70.

Winett, R. A. & Winkler, R. C. Current behavior modification in the classroom: Be still, be quiet, be docile. *Journal of Applied Behavior Analysis*, 1972, *5*, 499-504.

Wolpe, J. *The practice of behavior therapy*. New York: Pergamon, 1969.

## RECOMMENDED READINGS

Bergan, J. R. *Behavioral consultation*. Columbus, Ohio: Merrill, 1977.

This book is the best one available for helping you to acquire skills in conducting behavioral interviews with teachers and parents. In addition, it clearly demonstrates how to conduct functional analyses of student behaviors. Chapters deal with evaluation of behavioral programs and techniques for designing, implementing, and monitoring behavioral programs. Included are numerous checklists that the reader can use to evaluate his or her interviewing skills.

Hersen, M., & Bellack, A. S. *Behavioral assessment: A practical handbook*. New York: Pergamon, 1966.

Excellent overview of issues surrounding behavioral assessment, the historical development of assessment practices, and implementation and evaluation of behavioral programs. The two chapters on the assessment of childrens' behavioral deficits and excesses are invaluable for school psychologists working in the elementary schools. Specific guidelines are presented for conducting behavioral observations in classrooms and behavioral interviews with teachers and the clients themselves.

Ellis, A., & Grieger, R. *Handbook of rational emotive therapy*. New York: Springer, 1977.

This book will give school psychologists practical suggestions for assessing the cognitive and affective dimensions of student problems. It discusses both the multimodal assessment and treatment of emotional disturbance supplemented by case studies.